

STIC-Biot ch/ChemLib

CRFE

100926

8/13 03

From: Nickol, Gary  
Sent: Tuesday, August 12, 2003 10:48 AM  
To: STIC-Biotech/ChemLib  
Subject: 10/021660

RECEIVED MEJ

AUG 12 2003 8/13 43

8/14 171

(STIC) 8/15 135

8/18 76

8/19 61

8/20 6

Please search and interference search the following:

SEQ ID NO:41

Thanks,

Gary Nickol  
CM1, AU:1642  
Room 8D17, Mailbox 8E12  
(703) 305-7143

Point of Contact:  
Toby Port  
Technical Info. Specialist  
CM1 6A04  
703-308-3534

41-2456 na

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 8/13  
Date Completed: 8/20  
Searcher Prep/Review: DC  
Clerical: \_\_\_\_\_  
Online time: 12

TYPE OF SEARCH:  
NA Sequences: 1  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: CG  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2003, 11:41:37 ; Search time 8705 Seconds  
(without alignments)  
11542.107 Million cell updates/sec

Title: US-10-021-660-41

Perfect score: 2456

Sequence: 1 gcaagcacggaacaagctga.....aatgtggaacacccg 2456

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb.ba:\*

2: gb.htg:\*

3: gb.in:\*

4: gb.ov:\*

5: gb.pat:\*

6: gb.ph:\*

7: gb.pl:\*

8: gb.pr:\*

9: gb.ro:\*

10: gb.sts:\*

11: gb.un:\*

12: gb.vi:\*

13: gb.ba:\*

14: em.fun:\*

15: em.in:\*

16: em.mu:\*

17: em.ov:\*

18: em.ph:\*

19: em.pl:\*

20: em.ro:\*

21: em.sts:\*

22: em.un:\*

23: em.vi:\*

24: em.htg\_hum:\*

25: em.htg\_inv:\*

26: em\_htg\_mus:\*

27: em\_htg\_pln:\*

28: em\_htg\_rod:\*

29: em\_htg\_mam:\*

30: em\_htg\_vrt:\*

31: em\_sy:\*

32: em\_higo\_hum:\*

33: em\_higo\_mus:\*

34: em\_higo\_other:\*

35: em\_higo\_mus:\*

36: em\_higo\_mus:\*

37: em\_higo\_mus:\*

38: em\_higo\_mus:\*

39: em\_higo\_mus:\*

40: em\_higo\_mus:\*

41: em\_higo\_mus:\*

42: em\_higo\_mus:\*

43: em\_higo\_mus:\*

44: em\_higo\_mus:\*

45: em\_higo\_mus:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2456	100.0	2456	6	AR270913
2	2456	100.0	2456	9	HSBMXGENE
3	2424.4	98.7	2449	6	AX244661 Sequence B
4	2404.2	97.9	2604	9	AF045459 Homo sapi
5	2397.6	97.6	2500	6	AR179663 Sequence
6	2397.6	97.6	2500	6	AR274488 Sequence
7	2393.2	97.4	2553	9	BC016652 Homo sapi
8	1641.8	66.8	2609	10	MMU88091
9	1639.6	66.8	2369	10	AF012104
10	494.6	20.1	2468	6	AX178901
11	494.6	20.1	2468	10	MUSBTACDNA
12	494.6	20.1	2493	10	BC053392
13	493	20.1	2485	10	MUSBPK
14	493	20.1	2505	6	I25434
15	485	19.7	1365	6	AX467404 Sequence 1
16	485	19.7	1365	6	AX467404 Sequence
17	485	19.7	2560	9	AX178899 Sequence
18	480.2	19.6	2546	10	MUSEMBX
19	480.2	19.6	2582	6	AX244654 Sequence
20	447.6	18.2	2162	5	REU85659
21	440.4	17.9	15946	9	AC003669
22	440.4	17.9	19492	9	AC097625
23	439.4	17.9	2574	6	AR148881
24	439.4	17.9	2574	6	AR270952 Sequence
25	436	17.8	1971	5	AF535118
26	433.6	17.7	2574	10	S53716
27	433	17.6	2598	10	BC037071
28	430.2	17.5	2564	9	HMTYRKINA
29	429.8	17.5	2578	10	HMTYRKINA
30	421.2	17.1	2255	10	AF285881
31	409	16.7	1680	10	MMU19607
32	409	16.7	2204	10	MMU16145
33	409	16.7	2221	10	MUSRLK
34	409	16.7	2282	10	D83201
35	409	16.7	2342	10	MUSPTKRL18
36	409	16.7	2445	10	D83202
37	398.4	16.2	4224	10	MUSTLK
38	396.8	16.2	1878	6	AX695694
39	396.8	16.2	4231	10	MUSTYRKIN
40	396.8	16.2	4294	6	AX695693
41	396.8	16.2	4294	10	MUSTITKA
42	392.8	16.0	3650	9	HUMPTKA
43	392	16.0	2480	10	MUSEMTX
44	388	15.8	3593	6	AX244656
45	385.4	15.7	1368	6	AX467388 Sequence

# ALIGNMENTS

RESULT 1	AR270913	2456 bp	DNA	linear	PAT 10-APR-2003
LOCUS	AR270913	Sequence 1476 from patent US 6500938.			
DEFINITION	Sequence 1476 from patent US 6500938.				
ACCESSION	AR270913				
VERSION	AR270913.1	GI:29702147			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 2456)				
AUTHORS	Au-Young, J. and Seilhamer, J.J.				
TITLE	Composition for the detection of signaling pathway gene expression				
JOURNAL	Patent: US 6500938-A 1476 31-DEC-2002;				
FEATURES	Location/Qualifiers				

Pred. No. is the number of results predicted by chance to have a









QY	121	CTTTTGTGTTTACCAAAACAAACCTTCTCTACTATGAATATGACAAAATGAAAGGGC	180
DB	121	CTTTTGTGTTTACCAAAACAAACCTTCTCTACTATGAATATGACAAAATGAAAGGGC	180
QY	181	AGCAGAAAAGNATCCATTGAATTAAGAAAATCAGATGTGTGGAGAAATGTAATCTCGAG	240
DB	181	AGCAGAAAAGNATCCATTGAATTAAGAAAATCAGATGTGTGGAGAAATGTAATCTCGAG	240
QY	241	GAGCAGACCCCTGTAGAGAGACAGTACCATTTCAGATTGCTATAAAGATGGGCTTCTC	300
DB	241	GAGCAGACCCCTGTAGAGAGACAGTACCATTTCAGATTGCTATAAAGATGGGCTTCTC	300
QY	301	TATGTCTATGCATCAAAATGAAGAGCGGAAAGTCAAGTGTGCAAGCAATTAACAAAAGAG	360
DB	301	TATGTCTATGCATCAAAATGAAGAGCGGAAAGTCAAGTGTGCAAGCAATTAACAAAAGAG	360
QY	361	ATAAGGGGTAAACCCCACTGCTGTGTCAGTACCATTGCTGCTGTCGACGGGAAG	420
DB	361	ATAAGGGGTAAACCCCACTGCTGTGTCAGTACCATTGCTGCTGTCGACGGGAAG	420
QY	421	TTCCTGTGTTGCCAGCAGCTGTAAAGCAGCCCGAGGTACCCCTCTGGGAAGCATAT	480
DB	421	TTCCTGTGTTGCCAGCAGCTGTAAAGCAGCCCGAGGTACCCCTCTGGGAAGCATAT	480
QY	481	GCTAATCTGCATCTGCACTCAATGAAGAGAAACACAGAGTTCACACCTTCCAGACAGA	540
DB	481	GCTAATCTGCATCTGCACTCAATGAAGAGAAACACAGAGTTCACACCTTCCAGACAGA	540
QY	541	GTGCTGAAGATACCTCGGGCAGTTCCTGTTCTCAAAATGAGTACCACTTCTCAAGTACC	600
DB	541	GTGCTGAAGATACCTCGGGCAGTTCCTGTTCTCAAAATGAGTACCACTTCTCAAGTACC	600
QY	601	ACTCTAGCCCAATATGACAAAGCAATCAAGAAAATGAGTGGTCCAGCCACCATCTTCA	660
DB	601	ACTCTAGCCCAATATGACAAAGCAATCAAGAAAATGAGTGGTCCAGCCACCATCTTCA	660
QY	661	AGTACCAGTCTAGCGCAATATGACAGCAACTCAAGAAAATCTATGGTCCAGCCCAAC	720
DB	661	AGTACCAGTCTAGCGCAATATGACAGCAACTCAAGAAAATCTATGGTCCAGCCCAAC	720
QY	721	TTCACATGCAATATATCCAAAGGAAGACTTCCCTGACTGGTGGCAAGTAAAGAACTG	780
DB	721	TTCACATGCAATATATCCAAAGGAAGACTTCCCTGACTGGTGGCAAGTAAAGAACTG	780
QY	781	AAAAGTAGCAGCAGTGAAGATGTTGCAAGAGTAAACAAAAGAAAGAAATGTGAAT	840
DB	781	AAAAGTAGCAGCAGTGAAGATGTTGCAAGAGTAAACAAAAGAAAGAAATGTGAAT	840
QY	841	CACACCACTCAAAAGATTTCATGGGAATTCCTCGAGTCAAGTTCATCTGAAGAGAGAA	900
DB	841	CACACCACTCAAAAGATTTCATGGGAATTCCTCGAGTCAAGTTCATCTGAAGAGAGAA	900
QY	901	AACCTGGATGATTATGACTGGTGGTGAACATCTCCAGATCACAACTCTGAACAGTTA	960
DB	901	AACCTGGATGATTATGACTGGTGGTGAACATCTCCAGATCACAACTCTGAACAGTTA	960
QY	961	CTCAGCAAAAGGAAAAGAGAGCAATTTATGTTAGAAATTCAGCCAAAGTGGGAATG	1020
DB	961	CTCAGCAAAAGGAAAAGAGAGCAATTTATGTTAGAAATTCAGCCAAAGTGGGAATG	1020
QY	1021	TACACAGTGTCTTATTTAGTAAAGCTGTGAATGATAAAAGGAAGTCTCAACATATAC	1080
DB	1021	TACACAGTGTCTTATTTAGTAAAGCTGTGAATGATAAAAGGAAGTCTCAACATATAC	1080
QY	1081	CAGTGCATACAAATGCTGAGACAAATTTATACCTGGCAGAAAATCTACTGTTTGTATTC	1140
DB	1081	CAGTGCATACAAATGCTGAGACAAATTTATACCTGGCAGAAAATCTACTGTTTGTATTC	1140
QY	1141	ATTCCAAAGCTTATTCATATATCATCAACAATTCACAGAGCATGATCACAGGCTCCGC	1200
DB	1141	ATTCCAAAGCTTATTCATATATCATCAACAATTCACAGAGCATGATCACAGGCTCCGC	1200
QY	1201	CACCTGTGTCAACAAGGCCAACAGGTCCCGGACTCTGTGCTCCCTGGGAAATGGAATC	1260

DB	1201	CACCTGTGTCAACAAGGCCAACAGGTCCCGGACTCTGTGCTCCCTGGGAAATGGAATC	1260
QY	1261	TGGGAACCTGAAAAGAGAGAGATTACCTTTGTTCAAGAGAGCTGGGAAGTGGCCAGTTGGA	1320
DB	1261	TGGGAACCTGAAAAGAGAGAGATTACCTTTGTTCAAGAGAGCTGGGAAGTGGCCAGTTGGA	1320
QY	1321	GTGGTCCAGCTGGGCAAGTGGGAAGGGCAGTATGATGTTGCTGTTAAGATGATCAAGGAG	1380
DB	1321	GTGGTCCAGCTGGGCAAGTGGGAAGGGCAGTATGATGTTGCTGTTAAGATGATCAAGGAG	1380
QY	1381	GGCTCCATGTCAGAAAGATGAATTTCTTTTCAGGAGGCCAGACTATGATGAACACTCAGCCAT	1440
DB	1381	GGCTCCATGTCAGAAAGATGAATTTCTTTTCAGGAGGCCAGACTATGATGAACACTCAGCCAT	1440
QY	1441	CCCAAGCTGGTTAAATTTCTATGAGAGTGTGTTCAAGGAATACCCCATATACATAGTACT	1500
DB	1441	CCCAAGCTGGTTAAATTTCTATGAGAGTGTGTTCAAGGAATACCCCATATACATAGTACT	1500
QY	1501	GAAATATATGAAGCAATGGCTGCTGCTGTAATTTACCTGAGGAGTCAAGGAAAGGACTTGAA	1560
DB	1501	GAAATATATGAAGCAATGGCTGCTGCTGTAATTTACCTGAGGAGTCAAGGAAAGGACTTGAA	1560
QY	1561	CCTTCCAGCTCTTAGAAATGTCTACGATGTCTGTGAAGGCATGCGCTTCTTTGGAGAGT	1620
DB	1561	CCTTCCAGCTCTTAGAAATGTCTACGATGTCTGTGAAGGCATGCGCTTCTTTGGAGAGT	1620
QY	1621	CACCAATTCATACACCGGACTTGGCTGCTGCTGTAATTTACCTGAGGAGTCAAGGAAAGGACTTGAA	1680
DB	1621	CACCAATTCATACACCGGACTTGGCTGCTGCTGTAATTTACCTGAGGAGTCAAGGAAAGGACTTGAA	1680
QY	1681	GTGAAAGTATCTGACTTTGGAAATGACAAAGTATGTTCTTGTAGTACCAAGTATGTCAAGTCA	1740
DB	1681	GTGAAAGTATCTGACTTTGGAAATGACAAAGTATGTTCTTGTAGTACCAAGTATGTCAAGTCA	1740
QY	1741	GTGGAACAAAGTTCCTCAGTCAAGTGGTCCAGTCCAGAGGTGTTTCTTACATTTCAATATAC	1800
DB	1741	GTGGAACAAAGTTCCTCAGTCAAGTGGTCCAGTCCAGAGGTGTTTCTTACATTTCAATATAC	1800
QY	1801	AGCAGCAAGTCAGAGCTATGGGCAATTTGGGATCTCTGATGTTGGAGGTGTTCAAGCTGGGG	1860
DB	1801	AGCAGCAAGTCAGAGCTATGGGCAATTTGGGATCTCTGATGTTGGAGGTGTTCAAGCTGGGG	1860
QY	1861	AAGCAGCCCTATGACTTTGATGACAACTCCAGAGTGGTGTGGAAGTCTCCAGAGGCCAC	1920
DB	1861	AAGCAGCCCTATGACTTTGATGACAACTCCAGAGTGGTGTGGAAGTCTCCAGAGGCCAC	1920
QY	1921	AGGCTTTTACCGCCCTTCCAGTCCGACATCTACCAATCATGTACAGCTGCTGG	1980
DB	1921	AGGCTTTTACCGCCCTTCCAGTCCGACATCTACCAATCATGTACAGCTGCTGG	1980
QY	1981	CACGAGCTTCCAGAAAAGCTCCCAATTTACGAACTCTCTGTTCCATTTGAACCACTT	2040
DB	1981	CACGAGCTTCCAGAAAAGCTCCCAATTTACGAACTCTCTGTTCCATTTGAACCACTT	2040
QY	2041	CGGAAAAGCAAGCATTTGAAGAGAAATTTAGGAGTGTGATAGAATATAGATG	2100
DB	2041	CGGAAAAGCAAGCATTTGAAGAGAAATTTAGGAGTGTGATAGAATATAGATG	2100
QY	2101	CTGGCCAGCATTTTCATTTTAAAGGAAAGTGAAGGCAATAGTAAATTTAGTAGT	2160
DB	2101	CTGGCCAGCATTTTCATTTTAAAGGAAAGTGAAGGCAATAGTAAATTTAGTAGT	2160
QY	2161	TTTTAATAGTCTTCTCTCTATTGTCATTTAGAAATGAACAGGAGGAAACAAAG	2220
DB	2161	TTTTAATAGTCTTCTCTCTATTGTCATTTAGAAATGAACAGGAGGAAACAAAG	2220
QY	2221	ATTCCTTTGAAATTTAGATCAAAATTTAGTAAATTTCTTTTATGCTGCTCTGATATAACAC	2280
DB	2221	ATTCCTTTGAAATTTAGATCAAAATTTAGTAAATTTCTTTTATGCTGCTCTGATATAACAC	2280
QY	2281	TTTCCAGCTTATAGCAGAGCAGCATTTTCAGACTGCAATATAGAGACTGTGTTCTATGCT	2340

Db	2280	TTTCCAGCCTATAGCAGAGCACATTTTCAGACTGCAATATATAGAGACTGTGTTTCATGTGT	2339
Qy	2341	AAAGACTGAGCAGAACTGAAATAATTAATTTATTTGGATATTAATCTTTCTTTTATATATGTC	2400
Db	2340	AAAGACTGAGCAGAACTGAAATAATTAATTTATTTGGATATTAATCTTTCTTTTATATGTC	2399
Qy	2401	ATTGTGCACACAAATTAATATATCTACCAAGTACAGAAATGTGGAAA	2446
Db	2400	ATTGTGCACACAAATTAATATATCTACCAAGTACAGAAATGTGGACA	2445
RESULT 4			
AF045459			
LOCUS	AF045459	2604 bp	mRNA linear PRI 01-APR-1998
DEFINITION	Homo sapiens Etk/Bmx cytosolic tyrosine kinase mRNA, complete cds.		
ACCESSION	AF045459		
VERSION	AF045459.1 GI:3002962		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. (bases 1 to 2604)		
TITLE	Oiu.Y., Robinson,D., Pretlow,T. and Kung,H.J.		
JOURNAL	Etk/Bmx, a Tyrosine Kinase with a Pleckstrin-Homology Domain, Is an Effector of PI 3-Kinase and Involved in IL6 Induced Neuroendocrine Differentiation of Prostate Cancer Cells		
REFERENCE	Proc. Natl. Acad. Sci. U.S.A. (1998) In press		
AUTHORS	2. (bases 1 to 2604)		
TITLE	Oiu.Y., Robinson,D., Pretlow,T. and Kung,H.J.		
JOURNAL	Direct Submission		
FEATURES	Submitted (30-JAN-1998) Molecular Biology, Case Western Reserve University, 10900 Euclid Ave, Cleveland, OH 44106, USA		
source	Location/Qualifiers		
	1. .2604		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
CDS	/codon_start=1		
	/product="Etk/Bmx cytosolic tyrosine kinase"		
	/protein_id="AAC08966.1"		
	/db_xref="GI:3002963"		
	/translation="METTFELFKKKDAAANQWKNDDNMTKSLLELLKRSQKKKM SPNYKERLFVTKTSLSYEDXMDKRGSRKSEIEKKIRCEKVNLEBOTPYEROYP FOIVYDKGLLYVTAINEESRSQWLKALQKEIRGNPHLLVKYHSGFVDKGLCCQOQC KAATGCTLWEATLANLTAYNEKHRYTPDPRLVKIPRAVPVLMKDAPSSSTTLAODY NESKNGVGPSSPSSTSLAODSNSSKKIYGSOPNFNMOYIPREDPDMQWVKLKSS SEDVASSNQKERNVHTTSSKISWEPSSSEEBEENLDDYDFWAGNISRSQSEQLLR QKGLAFVVRNSQVMGTVSLSKAVNDKKGTVKHVYHVTNAENKLYLAENICFDS IKPLIHYHONSAGMTTRHPYSVTKANKVPDSVSLANGIWEKLRREITLLKEIGSQF GYVLQKWKQGDVAVYKMTKEGMSDEFFQAOTMKLSHPKLVKFYGVGSKKEYPIY IVTEYTSNGLNLRHSGKLEPSPOLLCEMDYDCGMAFLESHQPIHDLAARNCLV DRDLCKVSDGFQTRVLDQDYVSVYSGTKFPVKWSAPEVHFYKSSKSDVWAFGLIM WEVFSLCKOPDYLDNSQVVLKVSQGRHLRYRHLASDTIYIQIMYSCWHELPKRPITFQ QLLLSLEPLEKDKH"		
BASE COUNT	855 a	513 c	567 g 669 t
ORIGIN			
Query Match	97.9%;	Score 2404.2;	DB 9; Length 2604;
Best Local Similarity	99.5%;	Pred. No. 0;	
Matches 2422; Conservative	0; Mismatches	8; Indels	3; Gaps 1;
Qy	21	GACGGATGATATATGGATACAAAATCTATCTTAGAAGAACTCTCTCAAAAGATACACA	80
Db	175	GAAGATGATATATGGATACAAAATCTATCTTAGAAGAACTCTCTCAAAAGATACACA	234
Qy	81	GCAAAAGAGAAAATGTACCAAAATTAATACAAAGAACGGCTTTTGTGTTGACCAAAAC	140
Db	235	GCAAAAGAGAAAATGTACCAAAATTAATACAAAGAACGGCTTTTGTGTTGACCAAAAC	294
Qy	141	AAACCTTTCTACTATGAATATGACAAAAATGAAAAGGGCAGCAGAAAAAGGATCCATTGA	200
Db	295	AAACCTTTCTACTATGAATATGACAAAAATGAAAAGGGCAGCAGAAAAAGGATCCATTGA	354
Qy	201	AATTAAGAAAATTCAGATGTGTGGAGAAAGTAAATCTCGAGGAGCAGACGCCCTGTAGAGAG	260
Db	355	AATTAAGAAAATTCAGATGTGTGGAGAAAGTAAATCTCGAGGAGCAGACGCCCTGTAGAGAG	414
Qy	261	ACAGTACCCATTTTCAGATTTCTATATAAGATGGGCTTCTATGTCTATGTCATGCAATCAATGA	320
Db	415	ACAGTACCCATTTTCAGATTTCTATATAAGATGGGCTTCTATGTCTATGTCATGCAATCAATGA	474
Qy	321	AGAGAGCCGAAGTCTAGTGGTTGAAAGCATTACAAAAAGAGATAAGGGGTAAACCCACCT	380
Db	475	AGAGAGCCGAAGTCTAGTGGTTGAAAGCATTACAAAAAGAGATAAGGGGTAAACCCACCT	534
Qy	381	GCTGTCCTCAAGTACCAATAGTGGTTCCTTCTGTCGACGGGAAGTTCCTCTGTGTCGACAGAG	440
Db	535	GCTGTCCTCAAGTACCAATAGTGGTTCCTTCTGTCGACGGGAAGTTCCTCTGTGTCGACAGAG	594
Qy	441	CTGTAAAGCAGCCCCAGGATGTACCTCTCTGGGAAGCATATGCTAAATCTGCATACTCCAGT	500
Db	595	CTGTAAAGCAGCCCCAGGATGTACCTCTCTGGGAAGCATATGCTAAATCTGCATACTCCAGT	654
Qy	501	CAATGAAGAGAAACACAGAGTTCCCACTTCCAGACAGAGTGCCTCAAGATACCTCTCGGGC	560
Db	655	CAATGAAGAGAAACACAGAGTTCCCACTTCCAGACAGAGTGCCTCAAGATACCTCTCGGGC	714
Qy	561	AGTTCCTCTCTCAAAATGGATGCACCATCTTCAAGTACCACCTCTAGCCCAATATGACAA	620
Db	715	AGTTCCTCTCTCAAAATGGATGCACCATCTTCAAGTACCACCTCTAGCCCAATATGACAA	774
Qy	621	CGAATCAAGAAAAAATATGGCTCCAGCCACCACCTCTTCAAGTACCACCTCTAGCCCAAT	680
Db	775	CGAATCAAGAAAAAATATGGCTCCAGCCACCACCTCTTCAAGTACCACCTCTAGCCCAAT	834
Qy	681	TGACAGCAACTCAAGAAAAATCTATGGCTCCAGCCCAAACCTTCAAAATGCAGTATATTC	740
Db	835	TGACAGCAACTCAAGAAAAATCTATGGCTCCAGCCCAAACCTTCAAAATGCAGTATATTC	894
Qy	741	AAAGGAAGACTTCCCTGACTGTGTGGCAAGTAAGAAACCTGAAAAGTAGCAGCAGTGA	800
Db	895	AAAGGAAGACTTCCCTGACTGTGTGGCAAGTAAGAAACCTGAAAAGTAGCAGCAGTGA	954
Qy	801	AGATGTTGCAAGCAGTAACCAAAAGAAAGAAATGTGAATCACACCACCTCAAAAGATTTC	860
Db	955	AGATGTTGCAAGCAGTAACCAAAAGAAAGAAATGTGAATCACACCACCTCAAAAGATTTC	1014
Qy	861	ATGGGAATTCCTGAGTCAAGTTTCATCTCAAGAAAGAGAAAACCTGGATGATTGACTG	920
Db	1015	ATGGGAATTCCTGAGTCAAGTTTCATCTCAAGAAAGAGAAAACCTGGATGATTGACTG	1074
Qy	921	GTTTGTGTTAAACATCTCCAGATACAACTCTGAACAGTTACTCAGACAAAAGGAAAAGA	980
Db	1075	GTTTGTGTTAAACATCTCCAGATACAACTCTGAACAGTTACTCAGACAAAAGGAAAAGA	1134
Qy	981	AGGAGCATTTATGGTTAGAAATTCGAGCCCAAGTGGGAATGTACACAGTGTCTTATTAG	1040
Db	1135	AGGAGCATTTATGGTTAGAAATTCGAGCCCAAGTGGGAATGTACACAGTGTCTTATTAG	1194
Qy	1041	TAGGCTGTGAATGATATAAAAGGAAGTGTCAAAACATTACCAGTGCATACAAATGCTGA	1100
Db	1195	TAGGCTGTGAATGATATAAAAGGAAGTGTCAAAACATTACCAGTGCATACAAATGCTGA	1254
Qy	1101	GAACAAATTAATACCTGGCAGAAAACCTACCTGTTTGTGTTGATTCATTTCCAAAAGCTTATTCATTA	1160
Db	1255	GAACAAATTAATACCTGGCAGAAAACCTACCTGTTTGTGTTGATTCATTTCCAAAAGCTTATTCATTA	1314
Qy	1161	TCATCAACCAATTCACGAGGCGATGATCATACGGCTCCGCCACCCCTGTGTGTCACAAAGGC	1220
Db	1315	TCATCAACCAATTCACGAGGCGATGATCATACG - - -CCGCCACCCCTGTGTGTCACAAAGGC	1371
Qy	1221	CAACAAGGTCGCCGACTCTGTGTCTCTGGGAATGGAATCTGGGAATCTGGAAGTCAAAAGAGAA	1280
Db	1372	CAACAAGGTCGCCGACTCTGTGTCTCTGGCAATGGAATCTGGGAATCTGGAAGTCAAAAGAGAA	1431

```
QY 1281 GATTACCTGTTGAAGGAGCTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTG 1340
DB 1432 GATTACCTGTTGAAGGAGCTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTG 1491
QY 1341 GAAGGGGAGATGATGATGTTGCTGTTAAGATGATCAAGAGAGGCTCCATGTCAGAAGATGA 1400
DB 1492 GAAGGGGAGATGATGATGTTGCTGTTAAGATGATCAAGAGAGGCTCCATGTCAGAAGATGA 1551
QY 1401 ATTCTTTCAGGAGGCCAGACTATGATGAAGTCAAGCTCCAGCTGGTAAATTTCTA 1460
DB 1552 ATTCTTTCAGGAGGCCAGACTATGATGAAGTCAAGCTCCAGCTGGTAAATTTCTA 1611
QY 1461 TGGAGTGTGTTCAAGGAATACCCATATACATAGTACTGAATATATAAGCAATGGCTG 1520
DB 1612 TGGAGTGTGTTCAAGGAATACCCATATACATAGTACTGAATATATAAGCAATGGCTG 1671
QY 1521 CTTGCTGAATPACCTGAGGAGTACGGAAGAGACTTGAACCTTCCAGCTCTTAGAAT 1580
DB 1672 CTTGCTGAATPACCTGAGGAGTACGGAAGAGACTTGAACCTTCCAGCTCTTAGAAT 1731
QY 1581 GTGCTAGAGTCTGTGAGGATGCGCTTCTTGGAGAGTCAACCAATTCATACACCGGGA 1640
DB 1732 GTGCTAGAGTCTGTGAGGATGCGCTTCTTGGAGAGTCAACCAATTCATACACCGGGA 1791
QY 1641 CTTGCTGCTGCTAACTGCTTGGTGGACAGAGATCTCTGCTGAAAGTATCTGACTTTGG 1700
DB 1792 CTTGCTGCTGCTAACTGCTTGGTGGACAGAGATCTCTGCTGAAAGTATCTGACTTTGG 1851
QY 1701 AATGACAAGGTATGTTCTTGATACCCAGTATGTCAGTTTCAGTCCGGAACAAAGTTTCCAGT 1760
DB 1852 AATGACAAGGTATGTTCTTGATACCCAGTATGTCAGTTTCAGTCCGGAACAAAGTTTCCAGT 1911
QY 1761 CAAGTGTGCTAGCTCCAGAGTGTTCATTAATCAATACAGCAGCAGTCAAGTATG 1820
DB 1912 CAAGTGTGCTAGCTCCAGAGTGTTCATTAATCAATACAGCAGCAGTCAAGTATG 1971
QY 1821 GGCAATTTGGGATCTGATGTTGGAGGTGTTTCAGCTGGGAGCAGCCCTATGACTTGTA 1880
DB 1972 GGCAATTTGGGATCTGATGTTGGAGGTGTTTCAGCTGGGAGCAGCCCTATGACTTGTA 2031
QY 1881 TGCAAACTCCAGTGTGTTCTGAAGGTCTCCAGGGCCACAGGCTTTACCGGCCCCACCT 1940
DB 2032 TGCAAACTCCAGTGTGTTCTGAAGGTCTCCAGGGCCACAGGCTTTACCGGCCCCACCT 2091
QY 1941 GGCAATCGGACACATCTACAGATCATGTACAGCTGCTGGCAGCAGCTTCCAGAAAGCG 2000
DB 2092 GGCAATCGGACACATCTACAGATCATGTACAGCTGCTGGCAGCAGCTTCCAGAAAGCG 2151
QY 2001 TCCACATTTTCAGCAACTCTGCTTCCATTTGACCACTTCGGGAAAGAACAGCAGATG 2060
DB 2152 TCCACATTTTCAGCAACTCTGCTTCCATTTGACCACTTCGGGAAAGAACAGCAGATG 2211
QY 2061 AAGAAGAAATAGGAGTGTGATAAGAAATGAATATAGATGCTGCCAGCATTTCATTCA 2120
DB 2212 AAGAAGAAATAGGAGTGTGATAAGAAATGAATATAGATGCTGCCAGCATTTCATTCA 2271
QY 2121 TTTTAGGAAGTAGGAAGGCAATAGTAATTTTAGCTAGTGTGTTAAATGAGTGTCTCTGTA 2180
DB 2272 TTTTAGGAAGTAGGAAGGCAATAGTAATTTTAGCTAGTGTGTTAAATGAGTGTCTCTGTA 2331
QY 2181 TTGCTATATTTAGAAATGAACAGGAGGAGAAACAAAGATTCCTTGAATTTAGATC 2240
DB 2332 TTGCTATATTTAGAAATGAACAGGAGGAGAAACAAAGATTCCTTGAATTTAGATC 2391
QY 2241 AAATAGTAAATTTGTTTATGCTGCTCCTGTATATACATCTTCCAGCCTATAGCAGAAG 2300
DB 2392 AAATAGTAAATTTGTTTATGCTGCTCCTGTATATACATCTTCCAGCCTATAGCAGAAG 2451
QY 2301 CACATTTTCAGACTGCAATATAGAGAGTGTGTTTCATGTTGAAGACTTGAGCAGAACTGAA 2360
DB 2452 CACATTTTCAGACTGCAATATAGAGAGTGTGTTTCATGTTGAAGACTTGAGCAGAACTGAA 2511
```

```
QY 2361 AAATTACTTATTGGATATTCAATCTTTCTTATATGTCATGTCACACAAATTAATA 2420
DB 2512 AAATTACTTATTGGATATTCAATCTTTCTTATATGTCATGTCACACAAATTAATA 2571
QY 2421 TACTACCAAGTACAGAAATGTGGAAGAAAAA 2453
DB 2572 TACTACCAAGTACAGAAATGTGGAAGAAAAA 2604

RESULT 5
ARI79663
LOCUS ARI79663 2500 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 3 from patent US 6326469.
ACCESSION ARI79663
VERSION ARI79663.1 GI:2021218
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2500)
AUTHORS Ullrich,A.; Gishizky,M. and Sures,I.
TITLE Megakaryocytic protein tyrosine kinases
JOURNAL Patent: US 6326469-A 3 04-DEC-2001.
FEATURES
Location/Qualifiers
source 1..2500
BASE COUNT 819 a 505 c 553 g 623 t
ORIGIN

Query Match 97.6%; Score 2397.6; DB 6; Length 2500;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2417; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 22 ACGGATGATATATATGGATACAAAATCTATTCTAGAGAACTCTTCTCTCAAAAGATCACAG 81
DB 70 AAGGATGATATATGGATACAAAATCTATTCTAGAGAACTCTTCTCAAAAGATCACAG 129
QY 82 CAAAAGAGAAATATCTACCAAAATATTACAAGAACGGCTTTTGTGTTTGACCAAAACA 141
DB 130 CAAAAGAGAAATATCTACCAAAATATTACAAGAACGGCTTTTGTGTTTGACCAAAACA 189
QY 142 AACCTTCTCTACTATGAATATACAAAATGAAAAGGGGCGAGAGAAAGATCCATTGAA 201
DB 190 AACCTTCTCTACTATGAATATGACAAAATGAAAAGGGGCGAGAGAAAGATCCATTGAA 249
QY 202 ATTAAGAAATACAGATGTGTGAGAAAGTAAATCTCGAGGAGCAGACGCTGTAGAGAGA 261
DB 250 ATTAAGAAATACAGATGTGTGAGAAAGTAAATCTCGAGGAGCAGACGCTGTAGAGAGA 309
QY 262 CAGTACCCATTTTCAGATGTCTATAAAGATGGGCTTCTCTATGTCTATGTCATCAATGAA 321
DB 310 CAGTACCCATTTTCAGATGTCTATAAAGATGGGCTTCTCTATGTCTATGTCATCAATGAA 369
QY 322 GAGAGCGAAGTCAAGTGTGTTGAAAGCATTACAAAAGAGATAGGGGTAAACCCACCCTG 381
DB 370 GAGAGCGAAGTCAAGTGTGTTGAAAGCATTACAAAAGAGATAGGGGTAAACCCACCCTG 429
QY 382 CTGGTCAAGTACCATAGTGGGTCTTCTGTCGACGGGAAGTTCTCTGTTGCCAGCAGAGC 441
DB 430 CTGGTCAAGTACCATAGTGGGTCTTCTGTCGACGGGAAGTTCTCTGTTGCCAGCAGAGC 489
QY 442 TGTAAAGCAGCCCGCAGGATGTACCTCTGTGGGAAGCATATGCTAAATCTGCTACTGCGATC 501
DB 490 TGTAAAGCAGCCCGCAGGATGTACCTCTGTGGGAAGCATATGCTAAATCTGCTACTGCGATC 549
QY 502 AATGAAGAGAAACACAGAGTTCCCACTTCCAGACAGAGTCTGAAAGTACCTCGGGA 561
DB 550 AATGAAGAGAAACACAGAGTTCCCACTTCCAGACAGAGTCTGAAAGTACCTCGGGA 609
QY 562 GTTCTCTGTTCTCAAAATGGATGGACCACTTCTCAAGTACCACTTAGCCCAATATGACAA 621
DB 610 GTTCTCTGTTCTCAAAATGGATGGACCACTTCTCAAGTACCACTTAGCCCAATATGACAA 669
```

QY 622 GAATCAAGAAAAAAGTCTGCTCCAGCCACCATCTTCAAGTACCAGTCTAGCGCAATAT 681  
DB 670 GAATCAAGAAAAAAGTCTGCTCCAGCCACCATCTTCAAGTACCAGTCTAGCGCAATAT 729  
QY 682 GACAGCAACTCAAGAAAAAATCTATGGCTCCAGCCAAACTTCAACATGCAATGATATATCCCA 741  
DB 730 GACAGCAACTCAAGAAAAAATCTATGGCTCCAGCCAAACTTCAACATGCAATGATATATCCCA 789  
QY 742 AGGGAAGACTTCCCTGACTGCTGCAAGTAAAGAAAGTGAAGTACAGCAGCAGTGA 801  
DB 790 AGGGAAGACTTCCCTGACTGCTGCAAGTAAAGAAAGTGAAGTACAGCAGCAGTGA 849  
QY 802 GATGTTGCAAGCAGTAACCAAAAAAGAAAGTGTGAATCACACCACTCAAGATTTCA 861  
DB 850 GATGTTGCAAGCAGTAACCAAAAAAGAAAGTGTGAATCACACCACTCAAGATTTCA 909  
QY 862 TGGGAATTCCTGATGATCAAGTTCATCTGAAGAGGAAAAACCTGGATGATATGACTGG 921  
DB 910 TGGGAATTCCTGATGATCAAGTTCATCTGAAGAGGAAAAACCTGGATGATATGACTGG 969  
QY 922 TTTGCTGATACATCTCCAGATCAATCTCAACAGTTACTCAGACAAAAGGAAAGAA 981  
DB 970 TTTGCTGATACATCTCCAGATCAATCTCAACAGTTACTCAGACAAAAGGAAAGAA 1029  
QY 982 GGAGCATTTATGGTTAGAAATTCAGCAAGTGGGAATGTACAGATGCTCCTATTTAGT 1041  
DB 1030 GGAGCATTTATGGTTAGAAATTCAGCAAGTGGGAATGTACAGATGCTCCTATTTAGT 1089  
QY 1042 AAGGCTGTGAATGATAAAAAAGAAAGTGTCAACATTTACCAAGTGCATACAAATGCTGAG 1101  
DB 1090 AAGGCTGTGAATGATAAAAAAGAAAGTGTCAACATTTACCAAGTGCATACAAATGCTGAG 1149  
QY 1102 AACCAATTAATACCTGGCAGAAAACTACTGTTTGAATCCATTTCCAAAGCTTATTCATAT 1161  
DB 1150 AACCAATTAATACCTGGCAGAAAACTACTGTTTGAATCCATTTCCAAAGCTTATTCATAT 1209  
QY 1162 CATCAACAAATTCAGCAGGATGATCACAGCGCTCCGCCACCTGTGTCAACAAAGGCC 1221  
DB 1210 CATCAACAAATTCAGCAGGATGATCACAGCGCTCCGCCACCTGTGTCAACAAAGGCC 1269  
QY 1222 AACAGGTCCCGACTCTGTCTCCCTGGGAAATGGAATCTGGGAACCTGAAAGAGAAGAG 1281  
DB 1270 AACAGGTCCCGACTCTGTCTCCCTGGGAAATGGAATCTGGGAACCTGAAAGAGAAGAG 1329  
QY 1282 ATTACCTGTTGAAGGACTGGGAAGTGGCCAGTTGGAGTGGTCCAGCTGGGCAAGTGG 1341  
DB 1330 ATTACCTGTTGAAGGACTGGGAAGTGGCCAGTTGGAGTGGTCCAGCTGGGCAAGTGG 1389  
QY 1342 AAGGGCAGTATGATGTTGCTGTTAAGATGATCAAGGAGGCTCCATGTCAGAGATGAA 1401  
DB 1390 AAGGGCAGTATGATGTTGCTGTTAAGATGATCAAGGAGGCTCCATGTCAGAGATGAA 1449  
QY 1402 TTCTTTTCAGGAGGCCAGACTATGATGAACCTCAGCCATCCCAAGCTGGTTAAATTTCTAT 1461  
DB 1450 TTCTTTTCAGGAGGCCAGACTATGATGAACCTCAGCCATCCCAAGCTGGTTAAATTTCTAT 1509  
QY 1462 GGAGTGTGTTCAAGGAATATCCCATATACATATGACTGAATATATAAGCAATGGCTGC 1521  
DB 1510 GGAGTGTGTTCAAGGAATATCCCATATACATATGACTGAATATATAAGCAATGGCTGC 1569  
QY 1522 TTGCTGAATTAACCTGAGGAGTCAGGAAAGGACTTGAACCTTCCAGCTCTTGAAGATG 1581  
DB 1570 TTGCTGAATTAACCTGAGGAGTCAGGAAAGGACTTGAACCTTCCAGCTCTTGAAGATG 1629  
QY 1582 TGCTACGATGCTGTGAAGGATGGCCCTTTTGGAGAGTCAACCAATTCATACACGGGAC 1641  
DB 1630 TGCTACGATGCTGTGAAGGATGGCCCTTTTGGAGAGTCAACCAATTCATACACGGGAC 1689  
QY 1642 TTGGCTGCTGTAACCTGCTTGGTGACAGAGATCTCTGTGGAAGATATCTGACTTTGGA 1701  
DB 1690 TTGGCTGCTGTAACCTGCTTGGTGACAGAGATCTCTGTGGAAGATATCTGACTTTGGA 1749  
QY 1702 ATGACAAGGTATGTTCTTGATGACCAAGTATGTCAGTTTCAGTCGGNACAAAGTTTCCAGTC 1761

DB 1750 ATGACAAGGTATGTTCTTGATGACCAAGTATGTCAGTTTCAGTCGGAACAAAGTTTCCAGTC 1809  
QY 1762 AAGTGGTCAAGTCCAGAGGTGTTTCAATTAATACAGCAGCAAGTTCAGACCTATGG 1821  
DB 1810 AAGTGGTCAAGTCCAGAGGTGTTTCAATTAATACAGCAGCAAGTTCAGACCTATGG 1869  
QY 1822 GCATTTGGGATCCCTGATGTGGAGGTGTTTCAGCCCTGGGAGAGCAGCCCTATGACTGTGAT 1881  
DB 1870 GCATTTGGGATCCCTGATGTGGAGGTGTTTCAGCCCTGGGAGAGCAGCCCTATGACTGTGAT 1929  
QY 1882 GACAACCTCCAGGTGGTTCCTGAAAGTCTCCAGGCCCACAGGCTTTACGGCCCCACCTG 1941  
DB 1930 GACAACCTCCAGGTGGTTCCTGAAAGTCTCCAGGCCCACAGGCTTTACGGCCCCACCTG 1989  
QY 1942 GCATCGGACACATCTACCAGATCATGTACAGCTGCTGCGCAGAGCTTCCAGAAAGCCT 2001  
DB 1990 GCATCGGACACATCTACCAGATCATGTACAGCTGCTGCGCAGAGCTTCCAGAAAGCCT 2049  
QY 2002 CCCACATTTACAGCAACTCCTGCTTCCATTTGAACCACTTCGGGAAAAAGACAAAGATTGA 2061  
DB 2050 CCCACATTTACAGCAACTCCTGCTTCCATTTGAACCACTTCGGGAAAAAGACAAAGATTGA 2109  
QY 2062 AGAAGAAATTAGGAGTGTCTGATAAGAATGAATATAGATGCTGCGCAGCATTTTTCATTCAT 2121  
DB 2110 AGAAGAAATTAGGAGTGTCTGATAAGAATGAATATAGATGCTGCGCAGCATTTTTCATTCAT 2169  
QY 2122 TTTAAGGAAAGTAGAAGGCATTAATTTTAGCTAGTTAGTTTAAATAGTGTCTCTGTAT 2181  
DB 2170 TTTAAGGAAAGTAGAAGGCATTAATTTTAGCTAGTTTAAATAGTGTCTCTGTAT 2229  
QY 2182 TGCTCTATTTTAGAAATGAACAGGAGGAAACAAAGATTCCCTTGAATTTAGATCA 2241  
DB 2230 TGTCTATTTTAGAAATGAACAGGAGGAAACAAAGATTCCCTTGAATTTAGTCA 2289  
QY 2242 AATTAGTAATTTGTTTATGCTCTCTGATATAACACTTTTCCAGCCCTATAGCAGAAC 2301  
DB 2290 AATTAGTAATTTGTTTATGCTCTCTGATATAACACTTTTCCAGCCCTATAGCAGAAC 2348  
QY 2302 ACATTTTCAGACTGCAATATAGAGACTGTGTTCATGTGTAAAGACTGACGAGAACTGAAA 2361  
DB 2349 ACATTTTCAGACTGCAATATAGAGACTGTGTTCATGTGTAAAGACTGACGAGAACTGAAA 2408  
QY 2362 AATTACTTATTTGGATATTCTTTCTTTTATATTGTCATTTGTCACAAATTAATAT 2421  
DB 2409 AATTACTTATTTGGATATTCTTTCTTTTATATTGTCATTTGTCACAAATTAATAT 2468  
QY 2422 ACTACCAAGTACAGAAATGTGGAAGAAAAA 2453  
DB 2469 ACTACCAAGTACAGAAATGTGGAAGAAAAA 2500

## RESULT 6

AR274488 LOCUS AR274488 2500 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 3 from patent US 6506578.  
ACCESSION AR274488  
VERSION AR274488.1 GI:29706943  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2500)  
AUTHORS Ullrich,A., Gishizky,M. and Sures,I.G.  
TITLE Nucleotide encoding megakaryocytic protein tyrosine kinases  
JOURNAL Patent: US 6506578-A 3 14-JAN-2003;  
FEATURES  
source 1..2500  
BASE COUNT 819 a 505 c 553 g 623 t  
ORIGIN

Query Match 97.6%; Score 2397.6; DB 6; Length 2500;



Best Local Similarity 99.4%; Pred. No. 0; Matches 2417; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY	22	ACGGATGATATATGGATACAAATCTATTCTAGAGAACTCTCTCAAAAGATCAGAC	81
DB	70	AAGGATGATATATGGATACAAATCTATTCTAGAGAACTCTCTCAAAAGATCAGAC	129
QY	82	CAAAAGAGAAATCTCAAAATTAATACAAAGAACGGCTTTTGTGACAAAGAA	141
DB	130	CAAAAGAGAAATCTCAAAATTAATACAAAGAACGGCTTTTGTGACAAAGAA	189
QY	142	AACCTTTCCTACTATGATATGACAAATGAAAGGGGCGAGAGAAAGGATCCAAATGAA	201
DB	190	AACCTTTCCTACTATGATATGACAAATGAAAGGGGCGAGAGAAAGGATCCAAATGAA	249
QY	202	ATTAAGAAATCAGATGTGGAGAAAGTAATCTCGAGGAGCAGACGCTGTAGAGAGA	261
DB	250	ATTAAGAAATCAGATGTGGAGAAAGTAATCTCGAGGAGCAGACGCTGTAGAGAGA	309
QY	262	CAGTACCCATTTTCAGATTGTCTATAAAGATGGCTTCTATGCTATGCTATGCAATCAATGAA	321
DB	310	CAGTACCCATTTTCAGATTGTCTATAAAGATGGCTTCTATGCTATGCTATGCAATCAATGAA	369
QY	322	GAGAGCCGAAGTCAGTGGTTGAAAGCATTACAAAAGAGATAAGGGGTAAACCCCACTG	381
DB	370	GAGAGCCGAAGTCAGTGGTTGAAAGCATTACAAAAGAGATAAGGGGTAAACCCCACTG	429
QY	382	CTGGTCAAGTACATAGTGGGTTCTTCTGAGCGGAAAGTTCCTGTTGGCAGAGAGC	441
DB	430	CTGGTCAAGTACATAGTGGGTTCTTCTGAGCGGAAAGTTCCTGTTGGCAGAGAGC	489
QY	442	TGTAAGCAGCCAGGATGTACCTCTGGGAGCATATGCTAACTGCATACTGCAGTC	501
DB	490	TGTAAGCAGCCAGGATGTACCTCTGGGAGCATATGCTAACTGCATACTGCAGTC	549
QY	502	AATGAAGAGAAACACAGAGTTCCACCTTCCAGACAGAGTGCTGAAGATACCTCGGCA	561
DB	550	AATGAAGAGAAACACAGAGTTCCACCTTCCAGACAGAGTGCTGAAGATACCTCGGCA	609
QY	562	GTTCTGTTCTCAAAATGGATGCACCATCTTCAAGTACACACTAGCCCAATATGACAA	621
DB	610	GTTCTGTTCTCAAAATGGATGCACCATCTTCAAGTACACACTAGCCCAATATGACAA	669
QY	622	GAATCAAGAGAAATCTATGGCTCCAGCCACCATCTTCAAGTACACACTAGCCCAATAT	681
DB	670	GAATCAAGAGAAATCTATGGCTCCAGCCACCATCTTCAAGTACACACTAGCCCAATAT	729
QY	682	GACAGCACTCAAGAGAAATCTATGGCTCCAGCCCAACTTCAAGTACAGTATATTTCCA	741
DB	730	GACAGCACTCAAGAGAAATCTATGGCTCCAGCCCAACTTCAAGTACAGTATATTTCCA	789
QY	742	AGGAGAGATTCCTGACTGTGGCAATAGAGAACTGAAAGTACAGCAGCAGTGA	801
DB	790	AGGAGAGATTCCTGACTGTGGCAATAGAGAACTGAAAGTACAGCAGCAGTGA	849
QY	802	GATGTTCAAGCAGTAAACAAAGAGAAATGTGAATCACACCCTCAAGATTTCA	861
DB	850	GATGTTCAAGCAGTAAACAAAGAGAAATGTGAATCACACCCTCAAGATTTCA	909
QY	862	TGGGAATTCCTGAGTCAAGTTCTATGAAGAGAGAAACCTTGGATGATTATCACTGG	921
DB	910	TGGGAATTCCTGAGTCAAGTTCTATGAAGAGAGAAACCTTGGATGATTATCACTGG	969
QY	922	TTTCTGTTAACTCTCCAGATCAATCTGAACAGTTACTCAGACAAAGGAAAGAA	981
DB	970	TTTCTGTTAACTCTCCAGATCAATCTGAACAGTTACTCAGACAAAGGAAAGAA	1029
QY	982	GGACATTTATGGTTAGAAATCGAGCAAGTGGGAATGTACAGTGTCTTATTTAGT	1041
DB	1030	GGACATTTATGGTTAGAAATTCGAGCAAGTGGGAATGTACAGTGTCTTATTTAGT	1089
QY	1042	AAGSCTCTGAATGATAAAGAGAACTCTCAAACTTACACAGTGCATACAAATGCTGAG	1101

DB	1090	AAGGCTGTGAATGATAAAAAGGAACCTGTCAAACATTACACGTGCTACAAATGCTGAG	1149
QY	1102	AACAAATTTATACCTGGCAGAAAACTACTGTTTTGATTCCATTCCAAAGCTTATTCATTAT	1161
DB	1150	AACAAATTTATACCTGGCAGAAAACTACTGTTTTGATTCCATTCCAAAGCTTATTCATTAT	1209
QY	1162	CATCAACAATTCAGCAGCATGATCAGACGGCTCGGCCACCTGTGTCAACAAAGGCC	1221
DB	1210	CATCAACAATTCAGCAGCATGATCAGACGGCTCGGCCACCTGTGTCAACAAAGGCC	1269
QY	1222	AACAAGTCCCGACCTGTGTCCCTGGGAAATGGAATCTGGAACTGAAAGAGAAGAG	1281
DB	1270	AACAAGTCCCGACCTGTGTCCCTGGGAAATGGAATCTGGAACTGAAAGAGAAGAG	1329
QY	1282	ATTACCTTGTGAAGAGCTGGGAAGTGCCAGTTTGGAGTGTGTCCAGCTGGGCAAGTGG	1341
DB	1330	ATTACCTTGTGAAGAGCTGGGAAGTGCCAGTTTGGAGTGTGTCCAGCTGGGCAAGTGG	1389
QY	1342	AAGGGCAGTATGATGCTTTAAGATGATCAAGGAGGCTCCATGTCCAGAGATGAA	1401
DB	1390	AAGGGCAGTATGATGCTTTAAGATGATCAAGGAGGCTCCATGTCCAGAGATGAA	1449
QY	1402	TTCTTTTCAGAGGCCAGACTATGATCAAACTCAGCATCCCAAGCTGGTTAAATTTCTAT	1461
DB	1450	TTCTTTTCAGAGGCCAGACTATGATCAAACTCAGCATCCCAAGCTGGTTAAATTTCTAT	1509
QY	1462	GGAGTGTGTTCAAGGAATACCCCATATACATAGTACTGACTGAATATATAAGCAATGGCTGC	1521
DB	1510	GGAGTGTGTTCAAGGAATACCCCATATACATAGTACTGACTGAATATATAAGCAATGGCTGC	1569
QY	1522	TTGCTCAATACCTGAGGAGTACGCGAAAGGACTTGAACCTTCCAGCTTTCAGAAATG	1581
DB	1570	TTGCTCAATACCTGAGGAGTACGCGAAAGGACTTGAACCTTCCAGCTTTCAGAAATG	1629
QY	1582	TGCTACGATGCTGTGAAGGCTGGCTTCTTGGAGAGTCAACAAATTCATACACCGGAC	1641
DB	1630	TGCTACGATGCTGTGAAGGCTGGCTTCTTGGAGAGTCAACAAATTCATACACCGGAC	1689
QY	1642	TTGGCTGCTGTAACCTGTTGTGGACAGAGATCTCTGTGTGAAAGTATCTGACTTTGGA	1701
DB	1690	TTGGCTGCTGTAACCTGTTGTGGACAGAGATCTCTGTGTGAAAGTATCTGACTTTGGA	1749
QY	1702	ATGACAAGTATGTTCTTCTGATCAGCAGTATGCTCAGTTCAGTCCGGAACAAAGTTCCAGTC	1761
DB	1750	ATGACAAGTATGTTCTTCTGATCAGCAGTATGCTCAGTTCAGTCCGGAACAAAGTTCCAGTC	1809
QY	1762	AAGTGTCTAGCTCCAGAGTGTTCATTTCAATACAGCAGCAAGTCAGAGTATGG	1821
DB	1810	AAGTGTCTAGCTCCAGAGTGTTCATTTCAATACAGCAGCAAGTCAGAGTATGG	1869
QY	1822	GCATTTGGGATCCTGATGTGGAGGTGTTTACAGCTTGGGAAGCAGCCCTATGACTTGTAT	1881
DB	1870	GCATTTGGGATCCTGATGTGGAGGTGTTTACAGCTTGGGAAGCAGCCCTATGACTTGTAT	1929
QY	1882	GACAACCTCCAGGTGGTTCCTGAAGTCTCCAGGGCCACAGGCTTTACCGGGCCACCTG	1941
DB	1930	GACAACCTCCAGGTGGTTCCTGAAGTCTCCAGGGCCACAGGCTTTACCGGGCCACCTG	1989
QY	1942	GCATCGGACCACTCTACAGATCATGTACAGCTGTGGCAGCAGCTTCCAGAAAGCGT	2001
DB	1990	GCATCGGACCACTCTACAGATCATGTACAGCTGTGGCAGCAGCTTCCAGAAAGCGT	2049
QY	2002	CCACATTTACAGAACTCTCTTCCATTTGAACCTTCCGGAAGAAAGCAAGCATTTGA	2061
DB	2050	CCACATTTACAGAACTCTCTTCCATTTGAACCTTCCGGAAGAAAGCAAGCATTTGA	2109
QY	2062	AGAAGAAATAGGAGTGTGATGAAGTGAATATAGATGCTGCCAGCATTTTCATTTCAT	2121
DB	2110	AGAAGAAATAGGAGTGTGATGAAGTGAATATAGATGCTGCCAGCATTTTCATTTCAT	2169
QY	2122	TTTAAGGAAGTAGGAAGCATTAAGTAAATTTAGCTAGTTTTTAAGTGTCTCTGTAT	2181
DB	2170	TTTAAGGAAGTAGGAAGCATTAAGTAAATTTAGCTAGTTTTTAAGTGTCTCTGTAT	2229



Db	849	AGGAAGACTCCCTGACTGGTGGCAAGTAAGAAAACTGAAAAAGTAGCAGCAGCAGTGAA	908
Qy	802	GATGTTGCAAGCAGTAACCAAAAAGAAAGAAATGTGAATCACACCACCTCAAGATTTC	861
Db	909	GATGTTGCAAGCAGTAACCAAAAAGAAAGAAATGTGAATCACACCACCTCAAGATTTC	968
Qy	862	TGGGAATTCCTTGAGTCAAGTTCATCTCTGAAGAAGAGAAAACTCGATGATTATGACTGG	921
Db	969	TGGGAATTCCTTGAGTCAAGTTCATCTCTGAAGAAGAGAAAACTCGATGATTATGACTGG	1028
Qy	922	TTTGCTGTAACATCTCCAGATCACAATCTGAACAGTTACTCAGACAAAAGGAAAGAA	981
Db	1029	TTTGCTGTAACATCTCCAGATCACAATCTGAACAGTTACTCAGACAAAAGGAAAGAA	1088
Qy	982	GGAGCATTTATGGTTAGAAATTCGAGCCCAAGTGGGAATGTACACAGTCTCCTATTAGT	1041
Db	1089	GGAGCATTTATGGTTAGAAATTCGAGCCCAAGTGGGAATGTACACAGTCTCCTATTAGT	1148
Qy	1042	AAGGCTGTGAATGATAAAAAAGAACTGTCAAAATTTACCAGTGTCAATCAAAATGCTGAG	1101
Db	1149	AAGGCTGTGAATGATAAAAAAGAACTGTCAAAATTTACCAGTGTCAATCAAAATGCTGAG	1208
Qy	1102	AACAAATTTACCTGGCGAGAAATCTACTGTTTGTGATCCATTCCTCAAGCTTATTCATTAT	1161
Db	1209	AACAAATTTATACCTGGCGAGAAATCTACTGTTTGTGATCCATTCCTCAAGCTTATTCATTAT	1268
Qy	1162	CATCAACAAATTCAGAGGAGATCATCAGCGCTCCGCCACCTGTGTCAACAAGGCC	1221
Db	1269	CATCAACAAATTCAGAGGAGATCATCAGCGCTCCGCCACCTGTGTCAACAAGGCC	1328
Qy	1222	AACAAGTTCCTCGACTCTGTCTCCTGGGAAATGGAATCTGGGAATCAAAAAGAGAAAG	1281
Db	1329	AACAAGTTCCTCGACTCTGTCTCCTGGGAAATGGAATCTGGGAATCAAAAAGAGAAAG	1388
Qy	1282	ATTACCTTTGTAAGAGAGCTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGCGAAGTG	1341
Db	1389	ATTACCTTTGTAAGAGAGCTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGCGAAGTG	1448
Qy	1342	AAGGGCAGATGATGTGCTGTTAAGATGATCAAGGAGGCTCCATGTCAGAGATGAA	1401
Db	1449	AAGGGCAGATGATGTGCTGTTAAGATGATCAAGGAGGCTCCATGTCAGAGATGAA	1508
Qy	1402	TTCTTTTCAGGAGGCCAGACTATGATGAACCTCAGCCATCCAGCTGGTTAAATTCAT	1461
Db	1509	TTCTTTTCAGGAGGCCAGACTATGATGAACCTCAGCCATCCAGCTGGTTAAATTCAT	1568
Qy	1462	GGAGTGTGTTCAAAGGAATACCCCATATACATGACTGACTGAATATATAGCAATGGCTGC	1521
Db	1569	GGAGTGTGTTCAAAGGAATACCCCATATACATGACTGACTGAATATATAGCAATGGCTGC	1628
Qy	1522	TTGCTGAATTTACCTGAGAGTCAAGGAAAGGACTTGAACCTTCCAGCTCTTAGAAATG	1581
Db	1629	TTGCTGAATTTACCTGAGAGTCAAGGAAAGGACTTGAACCTTCCAGCTCTTAGAAATG	1688
Qy	1582	TGCTACCATGCTCTGAGGAGTCCCTCTCTGAGAGTCAACCAATTCATACACGGGAC	1641
Db	1689	TGCTACCATGCTCTGAGGAGTCCCTCTCTGAGAGTCAACCAATTCATACACGGGAC	1748
Qy	1642	TTGCTGCTGCTAACTGCTTGGTGGACAGAGATCTCTGTGTGAAGTATCTGACTTTGGA	1701
Db	1749	TTGGCTGCTGCTAACTGCTTGGTGGACAGAGATCTCTGTGTGAAGTATCTGACTTTGGA	1808
Qy	1702	ATGACAAAGGTATGTTCTTGATGACCAAGTATGCTAGTTTCAGTCGGAACAAAGTTTCCAGTC	1761
Db	1809	ATGACAAAGGTATGTTCTTGATGACCAAGTATGCTAGTTTCAGTCGGAACAAAGTTTCCAGTC	1868
Qy	1762	AAGTGGTCAAGTCAGAGGTTTTCATTATCTAATTAACAGCAAGTCAAGCTAGAGTATGG	1821
Db	1869	AAGTGGTCAAGTCAGAGGTTTTCATTATCTAATTAACAGCAAGTCAAGCTAGAGTATGG	1928
Qy	1822	GCATTTTGGGATCCGATGTGGGAGTGTTCAGCCCTGGGAGCAGCCCTATGACTTGTAT	1881
Db	1929	GCATTTTGGGATCCGATGTGGGAGTGTTCAGCCCTGGGAGCAGCCCTATGACTTGTAT	1988
Qy	1882	GACAACCTCCAGGTGGTCTTGAAGTCTCCAGGGCCACAGGCTTTTACCGGCCACCTG	1941
Db	1989	GACAACCTCCAGGTGGTCTTGAAGTCTCCAGGGCCACAGGCTTTTACCGGCCACCTG	2048
Qy	1942	GCATCGGACACCATCTACAGATCATGTACAGCTGCTGGCACGAGCTTCCAGAAAAGCGT	2001
Db	2049	GCATCGGACACCATCTACAGATCATGTACAGCTGCTGGCACGAGCTTCCAGAAAAGCGT	2108
Qy	2002	CCACATTTTCAGAACTCTCTGCTTCCATTTGAACCACTTCCGGGAAAAAGACAGCATTTGA	2061
Db	2109	CCACATTTTCAGAACTCTCTGCTTCCATTTGAACCACTTCCGGGAAAAAGACAGCATTTGA	2168
Qy	2062	ACAAGAAATTAGAGTGGCTGATAAAGATGAATATAGATGCTGCCAGCATTTTCATTCAT	2121
Db	2169	ACAAGAAATTAGAGTGGCTGATAAAGATGAATATAGATGCTGCCAGCATTTTCATTCAT	2228
Qy	2122	TTTAAAGAAAGTAGGAAGGCATAGTAAATTTAGCTAGTGTGTTTAAATAGTGTCTCTGAT	2181
Db	2229	TTTAAAGAAAGTAGGAAGGCATAGTAAATTTAGCTAGTGTGTTTAAATAGTGTCTCTGAT	2288
Qy	2182	TGCTCTATTATTAGAAATGAACAGCAGGAAACAAAGATTCCTTGAATTTAGATCA	2241
Db	2289	TGCTCTATTATTAGAAATGAACAGCAGGAAACAAAGATTCCTTGAATTTAGATCA	2348
Qy	2242	RAATTAATTTTGTGTTTATGCTGCTCCCTCATATAACACTTTTCCAGCCTATACAGAGC	2301
Db	2349	RAATTAATTTTGTGTTTATGCTGCTCCCTCATATAACACTTTTCCAGCCTATACAGAGC	2407
Qy	2302	ACATTTTCAGACTGCAATATAGAGACTGTGTTTCATGTGTAAGACTGAGCAGAACTGAAA	2361
Db	2408	ACATTTTCAGACTGCAATATAGAGACTGTGTTTCATGTGTAAGACTGAGCAGAACTGAAA	2467
Qy	2362	AATTAATTTTGGATATTCATTTCTTTTATATTTGTCATGTCACAACTTAAATAT	2421
Db	2468	AATTAATTTTGGATATTCATTTCTTTTATATTTATATTGTCATGTCACAACTTAAATAT	2527
Qy	2422	ACTACCAAGTACAGAAATGTGGAAA	2447
Db	2528	ACTACCAAGTAAAAA	2553
RESULT 8			
MMU88091			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Mus musculus (house mouse)			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1. 2609			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="C57Bl/6"			
/db_xref="taxon:10090"			
/tissue_type="bone marrow"			
142..2097			
/codon_start=1			
/product="protein tyrosine kinase Bmx"			
/protein_id="AAB4770.1"			
CDS			
1			
2609 bp			
mRNA			
linear			
ROD 19-FEB-1997			
complete cds.			



/db\_xref="GI:1843639";  
translation="MESKSIIEELLKKSOOKKMPNNYKERLFVLTKTSLSYEYD  
KMRGSRGSGTEIKIRCEKVNLEEDTPVEROYPOIYKDLGLIYVAVSESRQW  
LKAQKEIRGNPHLLIKVHSGFFVDGKFLCCQSCKAAPGCTLMEAYADLHIAISDEK  
HRAFTPERLLKIPRAVPVLDKADSSGAILPQIDSYSKSCGSPQNTNIRIIPREDC  
PDWQVRLKSEEDIDACSNOLERNIASHSTKMSWGFPESSSEENLHAYDWFAGN  
ISRSOQLLRQKQKGAFAFMYNSQMGMYTVSLFSAKNVNDKGTVRKHVHTWNAEK  
LYLAENTCFDSIPKLIHYHOHNSACMITRLRHPVSTRANKVPVSVLGGSGWELKREE  
ITLLKELNGOFGVQLOKQWGOYDVAKMKIKEGAMSEDEFQEOATNMKLSHPKLKLV  
FYGVCSKPIPIYIVTEXITNGCLLNYLKHSGKLESCOLLEMDYVCEGAFPLESHQF  
IHRDLARNCLVDSLSVKVDFGNWRYVLDQIVSVGTRFPVKNWSEPEVFFKYS  
SKSDWFAPILMWEVFSLGKQPIYLDNDSEVVVVKVSGHRLYRPLQASDTIYQIMYSC  
WHELPKRPFTFQOOLLSAIEPLREDDK\*  
BASE COUNT 812 a 503 c 637 g 657 t  
ORIGIN

Query Match 66.8%; Score 1641.8; DB 10; Length 2609;  
Best Local Similarity 82.3%; Pred. No. 0;  
Matches 2020; Conservative 0; Mismatches 337; Indels 96; Gaps 8;  
7 ACAGGAAAGCTGACGCGATGATAATATGATGACAAATCTATTCTAGAGAACTCTT 66  
115 ACAGAGGAGGCCAAATGGAGCACAATATGGAGAGTAATCCATCTCGAAGAACTGCTT 174  
67 CTCAAAAGATCACAGCAAAAGAGAAATGTACCAATAATATACAAAGACGGCTTTT 126  
175 CTCAAAAGTCACAGCAAAAGAGAAATGTACCAATAATATACAAAGACGGCTTTT 234  
127 GTTTTGACCAAAACAACTTCTCTACTATGATGACAAATATGACAAATGAAAGGGGACGAGA 186  
235 GTTCTAACCAAAACAACTTCTCTACTATGATACGACAAATGAAAGAGGAGGAGCAGA 294  
187 AAAGGATCCATGTAATTAAGAAATACAGATGTGTGGAGAAATGAAATCTCGAGAGGAG 246  
295 AAAGGATCAATAGAAATTAAGAAATCAGATGTGTGGAGAAATGAAATCTCGAGAGGAG 354  
247 ACGCCTGTAGAGACATACCCATTCAGATGTGTCTATAAGATGGGCTTCTCTATGTC 306  
355 ACACGGTGGAGACAGTACCCATTTTCAGATGTGTCTATAAGATGGGCTTCTCTATGTC 414  
307 TATCATCAATGAAGAGAGCGGAGTCAAGTGTGTGGAAGATTAAGATTAAGAGATTAAG 366  
415 TATCATCAATGAAGAGAGCGGAGTCAAGTGTGTGGAAGATTAAGATTAAGAGATTAAG 474  
367 GGTAAACCCCACTGCTGCTCAAGTACCATGATGCTGTGTGAGCGGGAAGTTCCTG 426  
475 GGCACCCCTCACTGCTGCTCAAGTATCATGATGCTGTGTGAGGATGGAAGTTCCTG 534  
427 TGTTCAGCAGAGCTGTAAAGCAGCCCGCAGATGTACCCCTCTGGGAGGATATGCTAAT 486  
535 TGTTCAGCAGAGCTGTAAAGCAGCCCGCAGATGTACCCCTCTGGGAGGATATGCTAAT 594  
487 CTGCATCTCAGTCAATGAGAGAGAAACACAGAGTTCACCTTCCAGACAGAGTGTG 546  
595 CTGCATCTCAGTCAATGAGAGAGAAACACAGAGTTCACCTTCCAGAGAGGCTACTG 654  
547 AAGTACCTCGGGAGTTCCTGCTTCAAAATGATGATGATGATGATGATGATGATGATGAT 606  
655 AAGATTCAGGAGGAGTTCCTGCTTCAAAATGATGATGATGATGATGATGATGATGATGAT 714  
607 GCCCAATATGACAGCAATCAAGAAACAACTATGCTCTCCAGCCACCACTTCTCAAGTACC 666  
715 CCC----- 717  
667 AGTCTAGCGCAATATGACAGCAACTCAAGAAATCTATGGCTCCAGCCAACTTCAAC 726  
718 -----CAATATGACAGTATTCAAAGAAAGTTGTGTTCCTCCAGCCAACTCAAC 768  
727 ATGCAATATATTCAGGAGGAGACTTCCCTGACTGTGGGCAAGTAAGAAACTGAGAAAGT 786  
769 ATACGCTATATTCAGGAGGAGACTTCCCTGACTGTGGGCAAGTAAGAAACTGAGAAAGT 825  
787 AGCAGCAGCAGTGAAGATGTTGCAAGCAGTAACCAAAAGAGAAAGAAAT---GTGAATCAC 843

826 -----AGCGAGGAGACATTGCATGCAGTCAACCAACTGGAAGAAATATCGCCTCAC 879  
844 ACCACCTCAAGATTTTCATGGGAATTCCTCGAGTCAAGTTCATCTCAAGAGAGGAAAC 903  
880 AGCAGCTCAAGATGTCATGGGATTCCTCGAGTCAAGTTCATCAAGAGAGGAGAAAT 939  
904 CTGGATGATATGACTGGTTCCTGCTGATACTCTCCAGATCAAACTCTGAACAGTACTC 963  
940 CTGCATGCTTATGACTGGTTCCTGGATATCTCCAGTCAAACTCTGAGCAGTACTC 999  
964 ACACAAAAGGGAAGAGAGCAATTTATGTTAGAAATTCGAGCCAAAGTGGGAATGTAC 1023  
1000 AGACAAAAGGGAAGAGAGCAATTTATGTTGCGAATTCAGCCAGGATGGAATGTAC 1059  
1024 ACAGTGTCTCTATTATTAGTAGAGGCTGTGAATGATATAAGAGGACTCTCAACATACCAC 1083  
1060 ACTGTGTCTCTATTATTAGTAGAGGCTGTGAATGATATAAGAGGAACTCTCAAGCATACCAC 1119  
1084 GTGCATACAAATGCTGAGAACAAATATATACCTGGCAGAAATACCTGTTTGTGATTCATT 1143  
1120 GTGCATACTAATGCTGAAATTAATATACCTGGCTGAAATACCTGCTTGTGATTCATT 1179  
1144 CAAAAGCTTATTTCATATCATCAACAAATTCAGCAGGATGATACACGGCTCCGCCAC 1203  
1180 CAAAAGCTCATTCATCTATCACAACATAATTCAGCAGGATGATACACGGCTCCGCCAC 1239  
1204 CCGTGTGTCAACAAAGGCAACAAAGTCCCGGCTGTGTCTGCTGCGAGTGAATGCTGG 1263  
1240 CCGTGTGTCAACAAAGGCAACAAAGTCCCGGCTGTGTCTGCTGCGAGTGAATGCTGG 1299  
1264 CAACCTGAAAGAGAGAGATTAACCTTGTGTAAGGAGCTGGGAGTGGCAGTGTGGAGTG 1323  
1300 GAACCTGAAAGAGAGAGATTAACCTTGTGTAAGGAGCTGGGCAATGGCCAGTGTGGAGTG 1359  
1324 GTCCAGCTGGGCAAGTGGAGGGGAGTATGATGTTGCTGTGAAGATGATCAAGAGGGC 1383  
1360 GTCCAGCTGGGCAAGTGGAGGGGAGTATGATGTTGCTGTGAAGATGATCAAGAGGGT 1419  
1384 TCCATGTCAAGAGATGAATTTCTTCAAGAGGCGCCAGACTATGATGAACCTCAGCCATCCC 1443  
1420 GCCATGTCAAGAGATGAATTTCTTCAAGAGGCGCCAGACTATGATGAACCTCAGCCATCCC 1479  
1444 AAGCTGGTAAATTTCTATGAGTGTGTTCAAAAGGAAATACCCATATACATAGTACTGAA 1503  
1480 AAGCTGGTGAAGTTCTACGAGTATGCTCAAGAAATACCCATCTATATAGTAACTGAG 1539  
1504 TATATAAGCAATGGCTGTGCTGATTAACCTGAGGAGTACACGGAAGAGACTTGAACCT 1563  
1540 TATATAACAAATGGTGTGCTTGTCTTAATACCTGAAGAGTCAATGGGAAGAGACTAGAAGT 1599  
1564 TCCAGCTCTTGAAGATGCTACGATGCTGTGAAGGAGTGGCTTCTTGGAGAGTCAAC 1623  
1600 TCCAGCTCTTGAAGATGCTATGATGCTGTGAAGGAGTGGCTTCTTGGAGAGCCAT 1659  
1624 CAATTATACACCGGAGTGGCTGCTGTAACCTGCTGAGGAGTGGTGGAGAGATCTCTGTG 1683  
1660 CAGTTTCATACATCGAGATTTGGCTGCTCGGAATTTGGTGGAGAGTATCTCTCTGTG 1719  
1684 AAGTATCTGACTTTGGAAATGACAGGATGTTCTTGTGATGACCACTATGTCAGTTCAGTC 1743  
1720 AAGTCTCAGACTTTGGAAATGACAGGATGTTCTTGTGATGACCACTATGTCAGTTCAGTA 1779  
1744 GGAACAAAGTTTCCAGTCAAGTGGTCCAGCTCCAGAGGTTTTCATTACTTCAATACAGC 1803  
1780 GGAACCAAGTTTCCAGTCAAGTGGTCCAGCTCCAGAGGTTTTCATTACTTCAATACAGC 1839  
1804 AGCAAGTCAGAGCTATGGCATTGGGATTCCTGATGTTGGAGGAGTGTTCAGCTCGGGAG 1863  
1840 AGCAAGTCGAGTATGGCATTGGGATTCCTGATGTTGGAGGAGTGTTCAGCTAGGAGAG 1899  
1864 CAGCCCTATGACTTGTATGACAACTCCCAAGTGTGTTCTGAAGTGTCTCCCAAGGCGCACAG 1923

Db	1900	CAGCCCTATGACTTATATGATACTCCGAGGTGCTGTGTGAAGGTCTCCCGAGGCCACAGA	1959
Qy	1924	CTTTACGGGCCCCACCTCCGATCCGACACACCATCTACAGATCATGTACAGTCTGTGGCAC	1983
Db	1960	CTCTACCGGCCCAACTGGCATCAGACACCATCTACAGATCATGTACAGTCTGTGGCAT	2019
Qy	1984	GAGTTCCAGAAAGCGTCCCAATTTTCAGAACTCTCTGCTTCCATTTGAACCACTTCGG	2043
Db	2020	GAGTTCCAGAAAGCGCCCACTTTTCAGAACTCTCTGCTTCCATTTGAACCACTTCGG	2079
Qy	2044	GAATAACAGCAAGCTTCAAGAGAGAAATTCAGAGTCTGATGAAGAAATATAGATGCTG	2103
Db	2080	GAACAAGTAAACATCAAGAGAGAAATTCAGAGTCTGATGAAGAAATATAGATGCTG	2139
Qy	2104	GCCAGCATTTTCATTCATTTTAAGGAAGTAGGAAGGCATTAAGTAAATTTTAGCTAGTTT	2163
Db	2140	ACTAACCATCTCATTTGTTGTAAGGGAAGTAGGAAGACTCAATCCATGTAGCTGCTCT	2199
Qy	2164	TAATAGTGTCTCTGATGTTGCTATATTTA-----GAATGAACAAGGCAGGA	2212
Db	2200	TAACAATGGTCTATGCTCTCTTATCTGTTATATATATACCAAAAAAAGCAAGGTAGGA	2259
Qy	2213	AACAAAGATTCCTCTCAAAATTAGATCAAAATAGTAAATTTGTTTATGCTGCTCTGA	2272
Db	2260	AACAAGATGCTCTTCAAAATTTAGGTCAAAATTTGTTGTTGCTGCTCTTAA	2318
Qy	2273	TATAACACTTTTCAGCCCTATAGCAGAGACATTTTCAGACTGCAATATATAGACATGCT	2332
Db	2319	GGT-ATACTTCCAAGACTATGGTAGAGTTTCATTTTATGATGCAATTTACAATTTGAC	2377
Qy	2333	TCATGTTAAAGACTGAGCAGAACTGAAAAATTTACATTTATGATGATTTTCATTTTCT--	2390
Db	2378	T-GTGATAAGGTAGAGCAGAAATTTGAAAGGTGATTTGTTGGGAATCCCTTTTAA	2436
Qy	2391	--TTATATTCTATTGTCACACAAATTAATATATACACCAAGTACAGAAATGT	2441
Db	2437	AAAAAATTTGGCTAGTGTGATAATTTAAATATCTTTTCAAGAACCGAAATGT	2489
LOCUS	AF012104	2369 bp	linear
DEFINITION	Mus musculus cytoplasmic tyrosine kinase (bmx) mRNA	ROD 09-OCT-1997	
ACCESSION	AF012104		
VERSION	AF012104.1	GI:2293557	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Exman, N., Lymbousseki, A., Vastrik, I., Sarvas, K., Kaipainen, A. and Alitalo, K.		
TITLE	Bmx tyrosine kinase is specifically expressed in the endocardium and the endothelium of large arteries		
JOURNAL	Circulation 96 (6), 1729-1732 (1997)		
MEDLINE	97462826		
PUBMED	9323053		
REFERENCE	2 (bases 1 to 2369)		
AUTHORS	Exman, N. and Alitalo, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-JUL-1997) Molecular/Cancer Biology Laboratory, Haartman Institute, P.O. B 21 (Haartmaninkatu 3), Helsinki 00014, Finland		
FEATURES	Location/Qualifiers		
source	1. .2369		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="129sv"		
	/db_xref="taxon:10090"		
gene	/dev_stage="12 day p.c. embryo"		
	1. .2369		
	/gene="bmx"		

QY	746	AAGACTTCCCTGACTGCTGGCAAGTAAGAAACTGAAAGTAGCAGACGACTGAAGTG	805
Db	655	AAGACTGCCCTGACTGCTGGCAAGTAAGAAACTGAAAGTAGCAGACGACTGAAGTG	705
QY	806	TTGCAAGCAGTAACCAAAAGAAAGAAAT---GTGAATCAACACACCTCAAAAGATTTCAT	862
Db	706	TTGCAAGCAGTAACCAAAAGAAAGAAATATCGCCTCTCAGACACCTCAAAAGATTTCAT	765
QY	863	GGGAATTCCTGAGTCAGTTCATCTGGAAGAGGAAACCTCGATGATTATGACTGGT	922
Db	766	GGGAATTCCTGAGTTCATCTGGAAGAGGAAAGAAATCTGCATGCTTATGACTGGT	825
QY	923	TTGCTGGTAAACATCTCAGATCAACAATCTGACAGTGTACTCAGACAAAGGAAAGAG	982
Db	826	TTGCTGGGAATATCTCAGAGTCAACAATCTGACAGTGTACTCAGACAAAGGAAAGAG	885
QY	983	GAGCATTATAGTTAGAAATTCGAGCAAGTGGGAATGTACACAGTGTCTTATTTAGTA	1042
Db	886	GAGCATTATAGTTAGAAATTCGAGCAAGTGGGAATGTACACAGTGTCTTATTTAGTA	945
QY	1043	AGGCTGTGAATGATAAAAGGAAGTGTCAACATATTACCAGTGCATACAAATGCTGAGA	1102
Db	946	AGGCTGTGAATGATAAAAGGAAGTGTCAACATATTACCAGTGCATACAAATGCTGAGA	1005
QY	1103	ACAAATTTATCTGCGCAAAACCTACTGTTTGTGATTCATTCCTCAAAAGCTTATTCATATC	1162
Db	1006	ATAAATTTATCTGCGTGAACCTACTGCTTTGTGATTCATTCCTCAAAAGCTTATTCATATC	1065
QY	1163	ATCAACACAATTCAGCAGGATGATCACACGGCTCCGCGCCACCTGTGTCAACAAAGGCCA	1222
Db	1066	ACCAACATAATTCAGCAGGATGATCACACGGCTCCGCGCCACCTGTGTCAACAAAGGCCA	1125
QY	1223	ACAAGTCCCGGACTGTGTCCCTGGGAATGGAATCTGGGAATCTGGAAGTCAAGAGAGA	1282
Db	1126	ACAAGTCCCGGACTGTGTCCCTGGGAATGGAATCTGGGAATCTGGAAGTCAAGAGAGA	1185
QY	1283	TTACTTGTGTAAGGAGTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGA	1342
Db	1186	TTACTTGTGTAAGGAGTGGGCAATGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGA	1245
QY	1343	AGGGCAGTATGATGTTGCTGTTAAGATGATCAAGAGGGCTCCATGTCAAGAGATGAAT	1402
Db	1246	AGGGCAATATGATGTTGCTGTTAAGATGATCAAGAGGGTGCATGTCAAGAGATGAAT	1305
QY	1403	TCCTTCAGGAGGCCAGACATGATGATAAATCAAGCCATCCCAAGCTGGTAAATCTATG	1462
Db	1306	TCCTTCAGGAGGCCAGACATGATGATAAATCAAGCCATCCCAAGCTGGTAAATCTATG	1365
QY	1463	GAGTGTGTTCAAGGAATACCCCATATACATAGTGAATGATATATAGCAATGCTGCT	1522
Db	1366	GAGTGTGTTCAAGGAATACCCCATCTATATAGTGAATGATATATAGCAATGCTGCT	1425
QY	1523	TGCTGAATTTACCTGAGGAGTCAACGAAAGGACTTGAACCTTCCAGCTCTTAGAAATGT	1582
Db	1426	TGCTGAATTTACCTGAGGAGTCAACGAAAGGACTTGAACCTTCCAGCTCTTAGAAATGT	1485
QY	1583	GCTACGATGCTGTGAAGGCATGCCCTTCTTGGAGAGTCAACCAATTCATACACCGGACT	1642
Db	1486	GTTATGATGCTGTGAAGGCATGCCCTTCTTGGAGAGCCATCAGTTCATACATCGAGATT	1545
QY	1643	TGGCTGTCTGAATGCTGTGTGGACAGAGATCTCTGTGTGAAGATATCTGACTTTGAA	1702
Db	1546	TGGCTGTCTGAATGCTGTGTGGACAGATCTCTGTGTGAAGATATCTGACTTTGAA	1605
QY	1703	TGACAAGTATGTTCTGATCAGCAGTATGCTGATCAGTCCAGTCCGAAAGTTCACATCA	1762
Db	1606	TGACGAGATATGCTCTTGTGATGATCAGTATGCTGATCAGTCCAGTCCGAAAGTTCACATCA	1665
QY	1763	AGTGGTCAGCTCCAGAGTGTTCATTTACTTCAATACAGCAGCAAGTCCAGCTATGCGG	1822
Db	1666	AGTGGTCAGCTCCAGAGTGTTCATTTACTTCAATACAGCAGCAAGTCCAGCTATGATGGG	1725
QY	1823	CATTGGGATCCTGATGCTGGGAGGTGTTGAGCTCGGGAAGCAGCCCTATGACTTGTATG	1882
Db	1726	CATTGGGATCCTGATGCTGGGAGGTGTTGAGCTAGGGAAGCAGCCCTATGACTTATATG	1785
QY	1883	ACAATCCCAAGTGTCTGAGGTCTCCAGGGCCACAGGCTTTACCGGCCCACTGG	1942
Db	1786	ATAACTCCGAGTGTGTTGAGGTCTCCAGGGCCACAGACTCTACCGGCCCACTGG	1845
QY	1943	CATCGGACACCATCTACAGATCATGTACAGCTGCTGGCAGCAGCTTCAGAAAAAGCGTC	2002
Db	1846	CATCAGACACCATCTACAGATCATGTACAGCTGCTGGCAGCAGCTTCAGAAAAAGCGTC	1905
QY	2003	CCACATTCACCAACTCCTGCTTCCATTGAACCACTTCGGGAAAAAGCAAGCAATGAA	2062
Db	1906	CCACATTCACCAACTCCTGCTGCCATGAGAACCACTTCGGGAAAAAGCAATGAA	1965
QY	2063	GAAGAAATTTAGGAGTGTGATAAGAAATGAATATAGATGCTGGCCAGCAATTTTCATTCAT	2122
Db	1966	GAGAAATTTAGGAGTGTGATAAGAAATGAATATAGATGCTGGCCAGCAATTTTCATTCAT	2025
QY	2123	TTAAGGAAGTAGGAAGCATAAAGTAATTTAGCTAGTGTAAATAGTGTCTCTCTATTT	2182
Db	2026	GTAGGGAAGTAGGAAGCATAAAGTAATTTAGCTAGTGTAAATAGTGTCTCTCTATTT	2085
QY	2183	GTCTATTATTATA-----GAAATCAACAGCAGCAGGAAACAAAGATTCCCTTGAA	2231
Db	2086	GTCTATTATTATA-----GAAATCAACAGCAGCAGGAAACAAAGATTCCCTTGAA	2145
QY	2232	ATTATAGTCAAAATAGTAATTTTGTGTTTATGCTGCTCCGATATACACTTTCCAGCCCTA	2291
Db	2146	ATTATAGTCAAAATAGTAATTTTGTGTTTATGCTGCTCCGATATACACTTTCCAGCCCTA	2203
QY	2292	TAGCAGAGCAGCAATTTTCCAGCTGCAATATAGATGCTGTTTCATGTGTAAAGACTGAGC	2351
Db	2204	TGGTAGAAGTTCATTTTATGATTCGAATTTTCAAAATTTGACT-GTGTATAAGGTAGAGC	2262
QY	2352	AGAACTGAAAAATTAATTTATGATTCATTTCTTTCTCT-----TTATATTGCTATTTGTA	2407
Db	2263	AGAACTGAAAAATTAATTTATGATTCATTTCTTTCTCT-----TTATATTGCTATTTGTA	2322
QY	2408	CAACAATTAATATATACCTACCAAGTACAGAAATGTGGAAAAAATAAATAA 2453	
Db	2323	TCGATAATTAATATATCTTTTCAAGACGGAATAAATAAATAAATAAATAA 2368	
RESULT 10			
AX718901			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Mus sp.			
Mus sp.			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
Patent: WO 0238797-A 3 16-MAY-2002;			
Bristol-Myers Squibb Company (US)			
Location/Qualifiers			
1..2468			
/organism="Mus sp."			
/mol_type="genomic DNA"			
/db_xref="taxon:10095"			
BASE COUNT			
713 a 559 c 602 g 594 t			
ORIGIN			
Query Match			
20.1%; Score 494.6; DB 6;			
Best Local Similarity 64.0%; Pred. No. 1.2e-110;			
Matches 746; Conservative 0; Mismatches 419; Indels 0; Gaps 0;			



1007 GCCAAGTGGGAATGTACACAGTGTCTTATTATTAGTAAGGCTGTAATGATAAAAAAGGAA 1066  
1061 GCAAGAGTGAATAATACACCGTGTCTGTGTGTGCTAAATCTACTGGGAGGCTCAAGGG 1120  
1067 CTGTCAACAGTTACCAAGTGTCAACAAATGCTGAGAACAAATATATACCTGGCAGAAACT 1126  
1121 TGATCGCCCATACGTTGTGTGTCCACGACAGAGCCAGTATTAACCTGGCTGAGAAC 1180  
1127 ACTGTTTGTGATCCATCCCAAGCTTATTCATATATCAACACAAATTCAGCAGGCATGA 1186  
1181 ACCTCTTCACACCATCCCTGAGCTCATTAATCACTACCAACAACTCTGCAGGCCATCA 1240  
1187 TCACAGGCTCCGCCACCTGTGTCTCAACAAAGGCCAACAGGTCGCCGACTCTGTGTGCC 1246  
1241 TATCCAGGCTGAATAATCTGTGTCTAAACAAACAAACAAACGCGCTTCTACTGCAGGCC 1300  
1247 TGGGAATGAATCTGGGAAGTCAAAAGAGAGAGATTAACCTGTGTGAAGAGCTGGGAA 1306  
1301 TGGGCTATGATCATCGGAAATGATCCAAAGGACCTCACCTTCTTGAAGAGAGCTGGGA 1360  
1307 GTGCCAGTGTGGAGTGTCCAGCTGGGCAAGTGGAAAGGCGAGTATGATGTGTGCTGTA 1366  
1361 CTGACAAATTCGGTGTGTGTAATATGGAAGTGGAGGGCCAAATATGATGTGGCCATCA 1420  
1367 AGATGATCAAGAGGCTCCATGTCAAGAGATGAATTTCTCAGAGGCGCCACACTATGA 1426  
1421 AGATGATCAGAGAGGTTCCATGTTCGAGGATGAATTCATTGAGAGGCCAAAGTCATGA 1480  
1427 TGAACCTACGCTATCCCAAGCTGTGTAAATTCATGAGAGTGTGTCAAGAGATACCCCA 1486  
1481 TGAATCTTCCCATGAGAAGCTGTGTGAGTGTATGTCGCTGCACCAACAAACAGGCCCA 1540  
1487 TATACATAGTCACTGAATATATAGCAATGCTGTCTGCTGCTGCTGCTGCTGCTGCTG 1546  
1541 TCTTCATCATCAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1600  
1547 GAAAGAGCTTGAACCTTCCAGCTCTTGAAGATGTCTAGTGTCTGCTGCTGCTGCTGCTG 1606  
1601 GCACCGCTTCCAGACACAGAGCTGCTGAGATGTCCAAAGATGTCTGTAAGCAATGG 1660  
1607 CTTCTTGGAGAGTACCAATTCATACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1666  
1661 AATACTTGGAGTGGAGAGCTGCTTCCAGAGACCTTGGCAGCTCGCAAACTGTTTGGTAA 1720  
1667 ACAGAGATCTCTGTGAAGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1726  
1721 AGATCAAGAGTGTGTAAGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1780  
1727 AGTATGTCAGTTCAGTCCGAACAAAGTTCCAGTCAAGTGTGCTAGCTCCAGAGGTGTTTC 1786  
1781 AGTACACCACTCTGTAGGCTCCAGTTTCCAGTCCGCTGCTCCACCAAGAGTGTGTTA 1840  
1787 ATTACTTCAATACAGCAAGTCAAGCTAGAGATGTGGGCAATTTGGGATCCCTGATGTGGAG 1846  
1841 TGTATACAAAGTTCAGCAGCAAAATCTGACATCTGGGCTTTTGGGGTTTAAATGTGGAGA 1900  
1847 TGTTCACCTTGGGAAGCAGCCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1906  
1901 TCTACTCCCTGGGAAGATCCGCTATGAGAGATTTACTACAGTGAGCAGCAGAGAACACA 1960  
1907 TCTCCAGGCGCACAGGCTTTACCGGCCACCTGGCAGCTGGACACCATCTACCAAGATCA 1966  
1961 TTGCTCAAGCTTACGCTCTACAGGCTCATCTGGCATCAGAGAGGTTATATACCATCA 2020  
1967 TGTACAGCTGCTGGCAGAGCTTCCAGAAAGCGTCCACATTTTCAACAACTCTCTGTCTT 2026  
2021 TGTACAGCTGCTGGCAGAGAAAGCAGATGAACGCTCTAGTTTCAAAATCTCTTTGAGTA 2080  
2027 CCATTGAACCACTCTGGGAAAAAGA 2051  
2081 ACATTCTAGATGTGATGGTGAAGA 2105

RESULT 12  
BC053392  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REMARK  
COMMENT

FEATURES  
source

BC053392 2493 bp mRNA linear ROD 09-JUN-2003  
Mus musculus Bruton agammaglobulinemia tyrosine kinase, mRNA (CDNA  
clone MGC:60523 IMAGE:30075322), complete cds.  
BC053392  
MGC  
BC053392.1 GI:31419801

Mus musculus  
(house mouse)

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2493)

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.F., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abrams, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Rouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
12477932

2 (bases 1 to 2493)

Strausberg, R.  
Direct Submission  
Submitted (02-JUN-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk  
Email: [cgapops-remail.nih.gov](mailto:cgapops-remail.nih.gov)  
Tissue Procurement: Drs. Dennis Taub, Dan Longo (NIA, USA),  
Jonathan Keller (NCI, USA)  
cDNA Library Preparation: Yulan Piao and Minoru Ko  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland:  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)  
Akhtar, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, J., Haghighi, P.,  
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 112 Row: 0 Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 7709993.  
Location/Qualifiers  
1. 2493  
/organism="Mus musculus"  
/mol\_type="mRNA"

Qy	1427	TGAACCTCAGCCATCCCACAAGTGGTTAAATTCTATGAGAGTGCTTCAAAGGAATACCCCA	1486
Db	1445	TGAATCTTTCCCATAGAAGCGGTGCAGTGTATGGCGTCTGCACCAACAACAGGCCCA	1504
Qy	1487	TATACATAGTGACTGAATATATAAGCAATGSGTCCTTGCTGAATTACCTCAGGAGTCACG	1546
Db	1505	TC TTTCATCATCCGAGTACATGGCTNANTGGCTGCGCTCTTTGAAC TACTCCTGAGGAGATGC	1564
Qy	1547	GAAAAGGACTTGAACCCTTCCCAGCTCTTAGAATATGCTACGATGCTGTGTAAGGCATGG	1606
Db	1565	GGCACCGCTTCCAGACACACGACGCTGCTTCAGATGTGCAAAAGATGCTGTGTAAGCAATGG	1624
Qy	1607	CCTTCTTGGAGAGTCCACCAATTATACACCGGACCTTGGCTGCTCGTAACTGCTTGGTGG	1666
Db	1625	AATACTTTGGAGTCGAAGCAGTCTCTTCACAGAGACCTGGCAGCTCGAAACTGTTTGGTAA	1684
Qy	1667	ACAGAGATCTCTGTGTGAAGTATCTGACPTTTGGAATGACAAGGTATGTTCTTGATGACC	1726
Db	1685	ACGATCAAGGAGTTGTGAAGTATCTGACPTTGGCCTGTCTAGGTATGCTCTTGATGATG	1744
Qy	1727	AGTATGTCAAGTTCAGTTCGGAAACAAAGTTTCCAGTCAAGTGGTCAGCTCCAGAGTGTTTT	1786
Db	1745	AGTACACCAAGCTCTGTAGGCTCCAAGTTTCCAGTCCGGTGGTCTCCACCAGAAAGTGGTTA	1804
Qy	1787	ATTACTTCAAAATACAGCAGCAAGTCAGACGTATGGCATTTTGGGATCCTCATCTGGGAGG	1846
Db	1805	TGTTATGACAAATTCAGCAGCAAAATCTGACATCTGGGCTTTTGGGGTTTTAATGTGGAGA	1864
Qy	1847	TGTTCAAGCCTGGGAAGCAGCCCTATGACTTGTATGACAACTCCCAGTGGTTCTTGAAGG	1906
Db	1865	CTTACTCCCTGGGAAGATCCCGCTATGAGAGATTTACTAACAGTGAGGCAGCAGAACACA	1924
Qy	1907	TC TCCAGGCCACAGGCTTTTACCGGCCCACTTGGCATCGGCACACCATCTACCAATCA	1966
Db	1925	TTGCTCAAGGCTTTACGCTCTCTACAGCGCTCATCTGCATCAGAGAGGGTATATACCATCA	1984
Qy	1967	TGTACAGCTCTGGCAGGAGCTTCCAGAAAGCGTCCACATTTTTCAGCAACTCCTGTCCTT	2026
Db	1985	TGTACAGCTGCTGGCACGAGAAACAGATGAACGCTCCTAGTTTCAAAATTTCTCTTGAGTA	2044
Qy	2027	CCATTGAACCACTTCGGGAAAAAGA	2051
Db	2045	ACATTCTAGATGTGATGGATGAAGA	2069

MUSBPBK	MUSBPBK	2485 bp	mRNA	linear	ROD 27-APR-1993
LOCUS					
DEFINITION	Mus musculus B cell cytoplasmic tyrosine kinase (BPK) mRNA, complete cds.				

ACCESSION	L08967
VERSION	L08967.1
KEYWORDS	GI:192233 tyrosine kinase.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 2485)  
REFERENCE

**AUTHORS** Tsukada, S., Saffran, D.C., Rawlings, D.J., Parolini, O., Allen, R.C., Kilsak, I.J., Kubagawa, H., Mohandas, T.K., Quan, S., Belmont, J.W., Cooper, M.D., Conway, M.E. and Witte, O.N.

**TITLE** Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia

**JOURNAL** Cell 72 (2), 279-290 (1993)

MEDLINE 93145329  
PUBMED 8425221

COMMENT  
Original source text: Mus musculus (strain C57BL/6 x DBA/2) pre-B cell cDNA to mRNA.

```
FEATURES
source
Location/Qualifiers
1. .2485
/organism="Mus musculus"
/mol type="mpna"
```

	Query Match.	20.1%;	Score 494.6;	DB 10;	Length 2493;
	Best Local Similarity	64.0%;	Pred. No. 1.2e-110;		
Matches	746; Conservative	0;	Mismatches 419;	Indels	0; Gaps 0;
QY	887	CTCAAGAAGAGGAAACCTGGATGATTATGACTGGTTTCTGTGTFACATCTCCAGATCAC	946		
Db	905	CTGAAGCTGAGGACTTCATAGAGATGTATCAGTGGTATTCCAAGACACATGACPCGAAGTC	964		
QY	947	AATCTGAACAGTTACTCAGACAAGGAAAAGGAGCATTTATTGGTTAGAAAAATTCGA	1006		
Db	965	AAGCTGAGCAACTGCTAAGCAAGAGGGGAAAAGAGGAGTTTCATTGTGCGGAGACTCCA	1024		
QY	1007	GCCAAGTGGGAATGTFACACAGTGTCTCTATTATTTAGTAGGCTGTGAATGATAAAAAAGGAA	1066		
Db	1025	GCAAAGCTGGAAATACACCGTGTCTGTGTTTGTCTAAATCTACTGGGGAGCCCTCAAGGGG	1084		
QY	1067	CTGTCAAAACANTACCAGTGCATACAAATCCTCAGAACACAAATTAACCTGGGCAGAAACT	1126		
Db	1085	TGATCGCCCATTAGTGTGTGTTCCAGGCCACAGAGCCAGTATTACCTGGCTGAGAAAC	1144		
QY	1127	ACTGTTTTGATTCCTTCCAAAGCTTATTTCATTATCATCAACACAATTCAGCAGGCGATGA	1186		
Db	1145	ACCTCTTTCAGCACCATCCCTGAGCTCATTTAACTACCATCAACACAACCTCTGCAGGSCCTCA	1204		
QY	1187	TCACAGGGCTCCGGCACCCGTGTGTCAACAAGGCCCAACAGGTCCCAGACTCTGTGTGCC	1246		
Db	1205	TATCCAGGCTGGAATATCCTGTGTCTAAACAAAACAAAACGGCGCTTCTTACTTCGAGGCC	1264		
QY	1247	TGGGAATGGAATCTCGGAACTCAAAAGAGAACAGATTACCTTGTTCAGGAGCTCGGAA	1306		
Db	1265	TGGGCTATGATCATGGGAATTTGATCCAAGGACCTCACCTTCTTGAAGGAGACTTGGGA	1324		
QY	1307	GTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGGAAAGGGCAGTATGATGTTGCTGTGA	1366		
Db	1325	CTGGACAATTCGGTGTGCGTGAATATGGGNAGTGGAGGGGCCAATATGATGTGGCCATCA	1384		
QY	1367	AGATGATCAAGGAGGCTCCANGTCAGAAGATGAATTCCTTCAGGAGGCCCAAGACTATGA	1426		
Db	1385	AGATGATCAGAAAGGTTCCATCTCGGAGGATCAATTCATTGAAGAAGGCCAAAGTCTCATGA	1444		





```
QY 887 CTGAAGAAGAGAAACCTGGATGATTATGACTGGTTGCTGGGTAAACATCTCCAGATCAC 946
Db 945 CTGAAGCTCGAGGATCCCATAGAGATGTATGAGTGGTATTTCCAAAGCACATGACTCGAAGTC 1004
QY 947 AATCTGAACAGCTTACTCAGACAAAGGCAAGAGGAGCATTATGCTTGAAGATTCGA 1006
Db 1005 AAGCTGAGCAACTGCTAAAGCAAGAGGGGAAAGAGGAGGTTTCAATGTCAGAGACTCA 1064
QY 1007 GCCAAGTGGGAATGTACACAGCTGCTCTTATTAGTAAGGCTGTGAATGATAAAAAAGGAA 1066
Db 1065 GCAAGCTGGAAATACACCGTGTCTGTGTTTCTGCTAAATCTACTGGGAGCTCAAGGG 1124
QY 1067 CTGTCAACATACCAGCTGCATACAAATGTGTGAACAATATATACCTGGCAGAAACT 1126
Db 1125 TGATCCGCGCTAGCTGTGTGTTCCAGCCACAGAGCCAGTATATACCTGGCTGAGAAC 1184
QY 1127 ACTGTTTGTATCCATTCCAAAGCTTATTCATATCATCAACAAATTCAGCAGCATGA 1186
Db 1185 ACCTCTCAGCACCATCCCTGAGCTCATTAATACCATCAACACANCTCTGAGGCCCTCA 1244
QY 1187 TCACACGGCTCCGACCTGTGTCAACAAAGGCCAACAAAGTCCCGACTCTGTGTCC 1246
Db 1245 TATCCAGGCTGAATATCTGTCTCTAAACAAACAAACAAACGCGCTTCTACTGCGGSC 1304
QY 1247 TGGGAATGGAATCTGGGAATGAAAGAGAGAGATTAACCTTGTGAAGGAGCTGGAA 1306
Db 1305 TGGCTATGGATCATGGAATATGATCCAAAGGACCTCACCTTCTTGAAGGAGCTGGGA 1364
QY 1307 GTGGCCAGTTTGGAGTGTCTGAGTGGGCAAGTGAAGGGCAGTATGATGTGTGTGA 1366
Db 1365 CTGGACAAATTCGGTGTGTAATATGGGAAGTGGAGGGCCATATGATGTGGCCATCA 1424
QY 1367 AGATGATCAAGAGGGCTCCATGTCTGAGAAGATGAATTTCTTCAGGAGCCCGAGATCA 1426
Db 1425 AGATGATCAGAAAGTTTCCATGTCTGGAGGATGAATTCATGGAAGAACCCAAAGTCATGA 1484
QY 1427 TGAATCTCAGCCATCCCAAGCTGGTTAAATCTATGAGAGTGTGTTCAAAGGAATACCCCA 1486
Db 1485 TGAATCTTCCCATGAGAAGTGTGTGATGGGCTGTGACCCAAACAAACCCCA 1544
QY 1487 TATACATAGTACGTAATATATAAGCAATGCTGCTTCTGAATTAACCTGAGGAGTCACG 1546
Db 1545 TCTTCATCATCCAGGATACATGCTTATGCTGCTTCTGAACTACCTGAGGAGATGC 1604
QY 1547 GAAAGGACTTGAACCTTCCAGCTCTTAGAAATGTGTACGATGTCTCTGAAGGCATGG 1606
Db 1605 GGCACCGCTCCAGACACAGCAGCTGCTTGAGATGTGCAAGATGTCTGTGACCAATGG 1664
QY 1607 CCTTCTTGGAGAGTACCAATTCATACACGGGACTTGGCTGCTGCTGCTTAACCTGTGTGG 1666
Db 1665 AATACTTGGAGTCAAGCAGTCTCTTACAGAGACCTGGCAGCTCGAAACTGTTTGGTAA 1724
QY 1667 ACAGAGATCTGTGTGAAGTATCTGACTTTGGATGACAGAGTATGTTCTTGTATGACC 1726
Db 1725 ACAGTCAAGGAGTGTGAAAGTATCTGACTTTGGCTGTCTAGTATGCTTCTGTATGATG 1784
QY 1727 AGTATGTCAGTTCAGTTCGGAACAAGTTTCCAGTCAAGTGTGCTGAGTCCAGAGGTGTTTC 1786
Db 1785 AGTACACAGCTCTGTAGGCTCCAGTTTCCAGTTCGCTGCTTCCACCGAAGTGTCTTA 1844
QY 1787 ATTACTTCAATACAGCAGCAAGTCAAGCTATGGGCAATTTGGGATCTCTGATGTGGAGG 1846
Db 1845 TGATACGAAGTTCAGCAGCAAACTGACATCTGGCTTTTGGGGTTTATATGTGGGAGA 1904
QY 1847 TGTTACGCTTGGGAAGAGCGCCCTATGACTTGTATGACAACTCCCGAGTGGTTCTGTAAGG 1906
Db 1905 TCTACTCCCTGGGAAGATGCGGTATGAGATTTACTTAACAGTGTGAGACAGCAGACA 1964
QY 1907 TCTCCAGGGCCACAGGCTTTACCGGCCCTACCTGGCATCCGACACCACTATACCATCA 1966
Db 1965 TTGCTCAAGGCTTACGTTCTACAGGCTCTATCTGGCATCAGAGGGGTATATACCATCA 2024
```

```
QY 1967 TGTACAGCTGCTGGCAGCAGCTTCCAGAAAAGGTCCTCCACATTTTCAGCAACTCTCTT 2026
Db 2025 TGTACAGCTGCTGGCAGCAGGAAAGCAGATGAACGCTCTAGTTTCAAAATTTCTCTGAGTA 2084
QY 2027 CCATTCAACCACTTCGGGAAAAAGA 2051
Db 2085 ACATTCTAGATGTGATGATGAAGA 2109

RESULT 15
AX467404
LOCUS AX467404 1365 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 22 from Patent WO0234899.
ACCESSION AX467404
VERSION AX467404.1 GI:21900613
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Sims,M.J. and Hayes,D.J.
TITLE Assay
JOURNAL Patent: WO 0234899-A 22 02-MAY-2002;
GLAXO GROUP LIMITED (GB)
FEATURES
source
BASE COUNT 395 a 301 c 344 g 325 t
ORIGIN
Query Match 19.7%; Score 485; DB 6; Length 1365;
Best Local Similarity 63.5%; Pred. No. 2.6e-108;
Matches 740; Conservative 0; Mismatches 425; Indels 0; Gaps 0;

QY 887 CTGAAGAAGAGAAACCTGGATGATTATGACTGGTTGCTGGTAAACATCTCCAGATCAC 946
Db 194 CTGAAGCGGAGACACTCCATAGAAATGTATGAGTGGTATTTCCAAACACATGACTCGGAGTC 253
QY 947 AATCTGACAGCTTACTCAGACAAAGGAAAGAGGAGCATTATGTTAGAAATTCGA 1006
Db 254 AGGCTCAGCAACTGTCTAAGCAAGAGGGGAAAGAGGAGGTTTTCATTTGTCAGAGACTCCA 313
QY 1007 GCCAAGTGGGAATGTACACAGTGTCTTATTTAGTAAGGCTGTGAATGATAAAAAAGGAA 1066
Db 314 GCAAAAGCTGGCAAAATATACAGTGTCTGTGTTTGTCTAAATCCACAGGGGACCTCAAGGGG 373
QY 1067 CTGTCAAAACATTTACCACCGTGCATACAAATGCTGAGAACAAATATATACCTGGCAGAAACT 1126
Db 374 TGATACGTCATTATGTTGTGTCTCCACACCTTCAGAGCCAGTATTTACCTGGCTGAGAGC 433
QY 1127 ACTGTTTGTATTCCTCAAGCTTATTCATATATCATCAACACAAATTCAGCAGGATGA 1186
Db 434 ACCTTTTCAGCACCACCTCCCTGAGCTCTTAACTACCATCAGCACAACCTCTGCAGGACTCA 493
QY 1187 TCACAGCGGTCCGCCACCTGTGTCAACAAAGGCCAACAAAGTCCCGGACTCTGTGTCC 1246
Db 494 TATCCAGGCTCAAAATATCAGTGTCTCTCAACAAAAGAAATGCACCTTCCACTGCAAGGCC 553
QY 1247 TGGGAAATGGAATCTGGGAACCTGAAAGAGAGAGATTAACCTTGTGTAAGGAGCTCGGAA 1306
Db 554 TGGGATACGGATCATGGGAAATGTATCCAAAGGACCTGACCTTCTTGAAGGAGCTGGGA 613
QY 1307 GTGGCCATTTGGAGTGTCTCCAGCTGGGCAACTGGAAGGGGAGTATGTTGTGCTTA 1366
Db 614 CTGGACAAATTTGGGCTAGTGAAGTATGGGAATGGAGAGGCCAGTACGAGCTGGCCATCA 673
QY 1367 AGATCAAGAGGGGCTCCATGTCTCAGAGATGAATTTCTTCAGGAGGCCAGACTATCA 1426
Db 674 AGATGATCAAGAGGGCTCCATGTCTGTAAGATGAATTCATTTGAAGAAAGCAAAATCATGA 733
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2003, 12:31:17 ; Search time 165 Seconds  
(without alignments)  
6569.921 Million cell updates/sec

Title: US-10-021-660-41

Perfect score: 2456

Sequence: 1 gcaagcagcggaacagctga.....aatgtgaaaaaaaaaacccg 2456

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Issued Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCrUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2456	100.0	2456	4	US-09-016-434-1476
2	2397.6	97.6	2500	4	US-08-426-509A-3
3	2397.6	97.6	2500	4	US-08-232-545-3
4	2397.6	97.6	2500	5	PCT-US95-05008-3
5	493	20.1	2505	1	US-08-391-615-1
6	439.4	17.9	2574	3	US-09-142-529-2
7	439.4	17.9	2574	4	US-10-045-428A-2
8	220.2	9.0	3623	1	US-08-306-691B-35
9	211.6	8.6	2647	4	PCT-US93-020132-77
10	211.6	8.6	2647	5	PCT-US93-06251-77
11	207.6	8.5	1611	1	US-08-391-615-7
12	207.6	8.5	1611	5	PCT-US93-00445-3
13	207.6	8.5	1611	5	PCT-US93-00445-3
14	204	8.3	1602	1	US-07-820-011A-1
15	204	8.3	1602	5	PCT-US93-00445-1
16	176.4	7.2	1491	2	US-09-006-675-1
17	176.4	7.2	1491	3	US-09-228-603A-1
18	173.6	7.1	1548	4	US-09-099-053-1
19	171	7.0	1467	4	US-09-579-182-2
20	168.2	6.8	780	2	US-09-006-675-7
21	168.2	6.8	780	3	US-09-228-603A-7
22	168	6.8	2770	4	US-08-426-509A-5
23	168	6.8	2770	4	US-08-232-545-5
24	168	6.8	2770	5	PCT-US95-05008-5
25	168	6.8	7607	1	US-08-222-616-19
26	168	6.8	7607	4	US-08-446-648-19
27	168	6.8	7607	5	PCT-US95-04228-19

Sequence 1, Appli  
Sequence 18, Appl  
Sequence 83, Appl  
Sequence 7, Appli  
Sequence 1, Appli  
Sequence 40, Appl  
Sequence 14, Appl  
Sequence 82, Appl  
Sequence 12, Appl  
Sequence 15, Appl  
Sequence 15, Appl  
Sequence 1452, Ap  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 1483, Ap  
Sequence 13, Appl  
Sequence 9, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-016-434-1476  
; Sequence 1476, Application US/09016434  
; Patent No. 6500938

#### GENERAL INFORMATION:

; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304

#### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

#### CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,434

#### FILING DATE: HEREWITH

#### CLASSIFICATION:

#### PRIOR APPLICATION DATA:

#### APPLICATION NUMBER:

#### FILING DATE:

#### CLASSIFICATION:

#### ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1476:

#### SEQUENCE CHARACTERISTICS:

LENGTH: 2456 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g951234

US-09-016-434-1476

Query Match

100.0%; Score 2456; DB 4; Length 2456;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GCAAGCAGGAAACAGCTGAGCGATGATAATATGGATACAAATCTATTCTAGAGAA	60
Db	1	GCAAGCAGGAAACAGCTGAGCGATGATAATATGGATACAAATCTATTCTAGAGAA	60
Qy	61	CTTCTCTCAAAAGATCACAGCAAAAGAGAAATGTCCACCAATAATTACAAAGACGG	120
Db	61	CTTCTCTCAAAAGATCACAGCAAAAGAGAAATGTCCACCAATAATTACAAAGACGG	120
Qy	121	CTTTTGTGTTGACCAAAACAACTTTCCTACTATGAATATGACAAATGAAAAGGGC	180
Db	121	CTTTTGTGTTGACCAAAACAACTTTCCTACTATGAATATGACAAATGAAAAGGGC	180
Qy	181	AGCAGAAAGGATCCATTGAAATTAAGAAATCAGATGCTGGAGAAAGTAAATCTCGAG	240
Db	181	AGCAGAAAGGATCCATTGAAATTAAGAAATCAGATGCTGGAGAAAGTAAATCTCGAG	240
Qy	241	GAGCAGACGCTGTAGAGACAGTACCCATTTTCAGATTGTCATAAAGATGGCTTCTC	300
Db	241	GAGCAGACGCTGTAGAGACAGTACCCATTTTCAGATTGTCATAAAGATGGCTTCTC	300
Qy	301	TATGCTATGATCAAAATGAAGAGCGCGAAGTCAAGTGGTGAAGCAATACAAAGAG	360
Db	301	TATGCTATGATCAAAATGAAGAGCGCGAAGTCAAGTGGTGAAGCAATACAAAGAG	360
Qy	361	ATAAGGGGTAAACCCACCTGCTGCTCAAGTACCATTAGTGGTCTTCGTCGAGCGAAG	420
Db	361	ATAAGGGGTAAACCCACCTGCTGCTCAAGTACCATTAGTGGTCTTCGTCGAGCGAAG	420
Qy	421	TTCCCTGTGTTGCCAGCAGAGCTGTAAAGCAGCCCCAGGATGTACCCCTCTGGGAAGCATAT	480
Db	421	TTCCCTGTGTTGCCAGCAGAGCTGTAAAGCAGCCCCAGGATGTACCCCTCTGGGAAGCATAT	480
Qy	481	GCTAACTGCTACTGACGATCAATGAGAGAAACACAGAGTTCCTCCAGACAGA	540
Db	481	GCTAACTGCTACTGACGATCAATGAGAGAAACACAGAGTTCCTCCAGACAGA	540
Qy	541	GTGCTGAAGATACCTCGGGCAGTTCCTGTTCTCAAAATGGAATGATGACCATCTCAAGTACC	600
Db	541	GTGCTGAAGATACCTCGGGCAGTTCCTGTTCTCAAAATGGAATGATGACCATCTCAAGTACC	600
Qy	601	ACTCTAGCCCAATATGACAAACGAAATCAAGAAACAACTATGGCTCCAGCCACCATCTTCA	660
Db	601	ACTCTAGCCCAATATGACAAACGAAATCAAGAAACAACTATGGCTCCAGCCACCATCTTCA	660
Qy	661	AGTACAGTCTAGGCAATATGACAGCAACTCAAGAAATCTATGGCTCCAGCCCAAC	720
Db	661	AGTACAGTCTAGGCAATATGACAGCAACTCAAGAAATCTATGGCTCCAGCCCAAC	720
Qy	721	TTCAACATGCAATATTTCAAGGAAGACTTCCCTGACTGGTGGCAAGTAAAGAACTG	780
Db	721	TTCAACATGCAATATTTCAAGGAAGACTTCCCTGACTGGTGGCAAGTAAAGAACTG	780
Qy	781	AAAAGTAGCAGCAGCTGAAGATGTTGCAAGCAGTAAACCAAAAGAAAGAAATGTGAAT	840
Db	781	AAAAGTAGCAGCAGCTGAAGATGTTGCAAGCAGTAAACCAAAAGAAAGAAATGTGAAT	840
Qy	841	CACACACCTCAAGATTTTATGGGAATTTCCCTGAGTCAAGTTCATCTGAAGAGAGAA	900
Db	841	CACACACCTCAAGATTTTATGGGAATTTCCCTGAGTCAAGTTCATCTGAAGAGAGAA	900
Qy	901	AACCTGATGATTTATGACTGTTGCTGGTAAACATCTCCAGATCACAAATCTGAACAGTTA	960
Db	901	AACCTGATGATTTATGACTGTTGCTGGTAAACATCTCCAGATCACAAATCTGAACAGTTA	960
Qy	961	CTCAGACAAAGGGAAGAGGAGCATTTATGTTAGAAATTCGAGCCAAAGTGGGAATG	1020
Db	961	CTCAGACAAAGGGAAGAGGAGCATTTATGTTAGAAATTCGAGCCAAAGTGGGAATG	1020
Qy	1021	TACACAGTGTCTTATTTAGTAAAGGCTGTGAATGATATAAAGAAAGCAATTTAC	1080
Db	1021	TACACAGTGTCTTATTTAGTAAAGGCTGTGAATGATATAAAGAAAGCAATTTAC	1080

Qy	1081	CACGTGCATACAAATGCTGAGAAACAAATTTATACCTGGCAGAAATCTACTGTTTGTATTC	1140
Db	1081	CACGTGCATACAAATGCTGAGAAACAAATTTATACCTGGCAGAAATCTACTGTTTGTATTC	1140
Qy	1141	ATTCCAAAGCTTATTTCATTATCATCAACAAATTCAGCAGGATGATCACAGGCTCCGC	1200
Db	1141	ATTCCAAAGCTTATTTCATTATCATCAACAAATTCAGCAGGATGATCACAGGCTCCGC	1200
Qy	1201	CACCTGTGTCACAAAGGCGCAAGGTCCTCCGACTCTGTCTCCCTGGGAAATGGAATC	1260
Db	1201	CACCTGTGTCACAAAGGCGCAAGGTCCTCCGACTCTGTCTCCCTGGGAAATGGAATC	1260
Qy	1261	TGGAACTGAAAGAGAGATTAACCTTGTGTAAGAGAGCTGGGAAGTGGCCAGTTTGA	1320
Db	1261	TGGAACTGAAAGAGAGATTAACCTTGTGTAAGAGAGCTGGGAAGTGGCCAGTTTGA	1320
Qy	1321	GTGCTCCAGCTGGGCAAGTGGAAAGGCGAGTATGATGTTGCTGTTAAAGATGATCAAGAG	1380
Db	1321	GTGCTCCAGCTGGGCAAGTGGAAAGGCGAGTATGATGTTGCTGTTAAAGATGATCAAGAG	1380
Qy	1381	GGCTCCATGTCAGAGATGAATTTCTTCAGGAGGCGCCAGACTATGATGAACTCAGCCAT	1440
Db	1381	GGCTCCATGTCAGAGATGAATTTCTTCAGGAGGCGCCAGACTATGATGAACTCAGCCAT	1440
Qy	1441	CCCAAGCTGGTAAATTTCTATGGAGTGTGTTCAAAGGAATACCCATATACATAGTACT	1500
Db	1441	CCCAAGCTGGTAAATTTCTATGGAGTGTGTTCAAAGGAATACCCATATACATAGTACT	1500
Qy	1501	GAATATATAAGCAATGCTGCTGATTAACCTGAGGAGTACGGAAGAGACTTGAA	1560
Db	1501	GAATATATAAGCAATGCTGCTGATTAACCTGAGGAGTACGGAAGAGACTTGAA	1560
Qy	1561	CTTCCCAGCTTTAGAAATGTCTACGATGCTCTGGAAGGATGSCCTCTTGGAGAGT	1620
Db	1561	CTTCCCAGCTTTAGAAATGTCTACGATGCTCTGGAAGGATGSCCTCTTGGAGAGT	1620
Qy	1621	CACCAATTCATACACCGGCACTTGGCTGCTGCTAACTGCTGGTGACAGAGATCTCTGT	1680
Db	1621	CACCAATTCATACACCGGCACTTGGCTGCTGCTAACTGCTGGTGACAGAGATCTCTGT	1680
Qy	1681	GTGAAAGTATCTGACTTTGGAAATGACAAAGTATGTTGATGACAGATGTCAGTTCA	1740
Db	1681	GTGAAAGTATCTGACTTTGGAAATGACAAAGTATGTTGATGACAGATGTCAGTTCA	1740
Qy	1741	GTCCGAACAAAGTTTCCAGTCAAGTGTCCAGCTCCAGAGGTTTTCATTACTTCAATAC	1800
Db	1741	GTCCGAACAAAGTTTCCAGTCAAGTGTCCAGCTCCAGAGGTTTTCATTACTTCAATAC	1800
Qy	1801	AGCAGCAAGTCAGAGCTATGGGCAATTTGGGATCCTGTGTTGGAGGTTTTCAGCCTGGG	1860
Db	1801	AGCAGCAAGTCAGAGCTATGGGCAATTTGGGATCCTGTGTTGGAGGTTTTCAGCCTGGG	1860
Qy	1861	AAGCAGGCTTATGACTTTGTATGACAACTCCAGGTTGTTCTGAAGGTTCCAGGGCCAC	1920
Db	1861	AAGCAGGCTTATGACTTTGTATGACAACTCCAGGTTGTTCTGAAGGTTCCAGGGCCAC	1920
Qy	1921	AGGCTTTACCGGCCCCACCTGSCATCGGACACCATCTACCAGATGATGATGAGTCTGCG	1980
Db	1921	AGGCTTTACCGGCCCCACCTGSCATCGGACACCATCTACCAGATGATGATGAGTCTGCG	1980
Qy	1981	CACGAGCTTCCAGAAAGCGTCCCACTTTAGCAACTCCCTCTCTTCCATTGAACACATT	2040
Db	1981	CACGAGCTTCCAGAAAGCGTCCCACTTTAGCAACTCCCTCTCTTCCATTGAACACATT	2040
Qy	2041	CGGAAAGAGCAAGCATTTGAAGAGAAATTAGGAGTGTGATAGAATGAATATAGATG	2100
Db	2041	CGGAAAGAGCAAGCATTTGAAGAGAAATTAGGAGTGTGATAGAATGAATATAGATG	2100
Qy	2101	CTGGCCAGCATTTTCATTTCATTTTAAAGAAAGTAGAGGCAATAGTAAATTTAGCTAGT	2160
Db	2101	CTGGCCAGCATTTTCATTTCATTTTAAAGAAAGTAGAGGCAATAGTAAATTTAGCTAGT	2160

QY 2161 TTTTAAATAGTCTCTCTGTTATTTCTATTATTTAGAAATGAACAAGCGGAGAAACAAAG 2220  
|||||  
Db 2161 TTTTAAATAGTCTCTCTGTTATTTCTATTATTTAGAAATGAACAAGCGGAGAAACAAAG 2220  
|||||  
QY 2221 ATTCCCTTGAATTTAGATCAAAATAGTAATTTTCTTTTATGCTGCTCTCTGATATAACAC 2280  
|||||  
Db 2221 ATTCCCTTGAATTTAGATCAAAATAGTAATTTTCTTTTATGCTGCTCTCTGATATAACAC 2280  
|||||  
QY 2281 TTTCCAGCTATAGCAGAGACACATTTTCAGACTGCCATATATAGAGACTGTGTCAATGTC 2340  
|||||  
Db 2281 TTTCCAGCTATAGCAGAGACACATTTTCAGACTGCCATATATAGAGACTGTGTCAATGTC 2340  
|||||  
QY 2341 AAAGACTGAGCAGAACTGAAATTTACTATTGGATATTTCTTTCTTTATATTTGTC 2400  
|||||  
Db 2341 AAAGACTGAGCAGAACTGAAATTTACTATTGGATATTTCTTTCTTTATATTTGTC 2400  
|||||  
QY 2401 ATTGTACACAAATTAATATATACCAAGTACAGAAATGTGGGAAAAAACAACCG 2456  
|||||  
Db 2401 ATTGTACACAAATTAATATATACCAAGTACAGAAATGTGGGAAAAAACAACCG 2456  
|||||

## RESULT 2

US-08-426-509A-3

; Sequence 3, Application US/08426509A  
; Patent No. 6326469

## GENERAL INFORMATION:

; APPLICANT: Ullrich, Axel  
; APPLICANT: Gishizky, Mikhail  
; APPLICANT: Sures, Irman G.  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
; TITLE OF INVENTION: TYROSINE KINASES  
; NUMBER OF SEQUENCES: 21

## CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York,  
; STATE: NY  
; COUNTRY: USA

; ZIP: 10036-2711

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0

## CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/426,509A  
; FILING DATE: 21-APR-1995

## CLASSIFICATION:

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/232,545

## FILING DATE:

## ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-0074-999

## TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-9741

; TELEX: 66141 PENNIE

## INFORMATION FOR SEQ ID NO: 3:

## SEQUENCE CHARACTERISTICS:

; LENGTH: 2500 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

US-08-426-509A-3

## Query Match

Best Local Similarity 97.6%; Score 2397.6; DB 4; Length 2500;

Matches 2417; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 22 ACGGATGATATATGGATACAAATCTATTCTAGAGAACTCTCTCTCAAAAGATCACAG 81  
| |||||

Db 70 AAGGATGATATATGGATACAAATCTATTCTAGAGAACTCTCTCTCAAAAGATCACAG 129  
QY 82 CAAAAGAGAAAATGTCACCAATAATATTTACAAAGAACGGCTTTTGTGTTTGTGACCAAAACA 141  
|||||  
Db 130 CAAAAGAGAAAATGTCACCAATAATATTTACAAAGAACGGCTTTTGTGTTTGTGACCAAAACA 189  
|||||  
QY 142 AACCTTTTCTACTATGCAATATGACAAATGCAAAAGGGGAGAGAGAGAGATCCATTTGAA 201  
|||||  
Db 190 AACCTTTTCTACTATGCAATATGACAAATGCAAAAGGGGAGAGAGAGAGATCCATTTGAA 249  
|||||  
QY 202 ATTAAGAAAATCAGATGTGTGGAGAAAGTAAATCTCGAGGAGAGAGAGAGAGAGAGAGAG 261  
|||||  
Db 250 ATTAAGAAAATCAGATGTGTGGAGAAAGTAAATCTCGAGGAGAGAGAGAGAGAGAGAGAG 309  
|||||  
QY 262 CAGTACCCATTTTCAGATTTCTATTAAGATGGCTTCTCTATGCTATGATCAATATGAA 321  
|||||  
Db 310 CAGTACCCATTTTCAGATTTCTATTAAGATGGCTTCTCTATGCTATGATCAATATGAA 369  
|||||  
QY 322 GAGAGCCGAAGTCACTGTTGAAAGCATTTACAAAAGAGAGATAGGGGTAAACCCCACTG 381  
|||||  
Db 370 GAGAGCCGAAGTCACTGTTGAAAGCATTTACAAAAGAGAGATAGGGGTAAACCCCACTG 429  
|||||  
QY 382 CTGGTCAAGTACCATAGTGGGTTCTTCGTGGAGCGGGAAGTTCCTGTGTGCCAGCAGAGC 441  
|||||  
Db 430 CTGGTCAAGTACCATAGTGGGTTCTTCGTGGAGCGGGAAGTTCCTGTGTGCCAGCAGAGC 489  
|||||  
QY 442 TGTAAAGCAGCCCCAGGATGTACCCCTCTGGGAAGCATATGCTTAATCTGCATCTGCGAGTC 501  
|||||  
Db 490 TGTAAAGCAGCCCCAGGATGTACCCCTCTGGGAAGCATATGCTTAATCTGCATCTGCGAGTC 549  
|||||  
QY 502 AATGAAGAGAAAACACAGAGTTCACCCTTCCCAAGAGAGAGTGTGAAGATACCTTCGGGCA 561  
|||||  
Db 550 AATGAAGAGAAAACACAGAGTTCACCCTTCCCAAGAGAGAGTGTGAAGATACCTTCGGGCA 609  
|||||  
QY 562 GTTCTGTGTTCTCAAAATGGATGACCATCTTCAAGTACCCTCTAGCCCAATATGCAAC 621  
|||||  
Db 610 GTTCTGTGTTCTCAAAATGGATGACCATCTTCAAGTACCCTCTAGCCCAATATGCAAC 669  
|||||  
QY 622 GAATCAAGAAAACATATGCTGCCAGCCAGCCATCTTCAAGTACCCTCTAGCCCAATAT 681  
|||||  
Db 670 GAATCAAGAAAACATATGCTGCCAGCCAGCCATCTTCAAGTACCCTCTAGCCCAATAT 729  
|||||  
QY 682 GACAGCACTCAAGAAAATCTATGCTCTCCAGCCAAACTTCAACATGCAAGTATATTTCCA 741  
|||||  
Db 730 GACAGCACTCAAGAAAATCTATGCTCTCCAGCCAAACTTCAACATGCAAGTATATTTCCA 789  
|||||  
QY 742 AGGGAAGACTTCCCTGACTGGTGGCAAGTAAGAAAACCTGAAAAGTAGCAGCAGAGTCAA 801  
|||||  
Db 790 AGGGAAGACTTCCCTGACTGGTGGCAAGTAAGAAAACCTGAAAAGTAGCAGCAGAGTCAA 849  
|||||  
QY 802 GATGTTGCAAGCAGTAACCAAAAAGAAAAGAAATGTGAATCACACACCCTCAAGATTTCA 861  
|||||  
Db 850 GATGTTGCAAGCAGTAACCAAAAAGAAAAGAAATGTGAATCACACACCCTCAAGATTTCA 909  
|||||  
QY 862 TGGGAATTCCTGAGTCAAGTTCATCTGAAGAGAGAGAAAACCTGATGATATGACTGG 921  
|||||  
Db 910 TGGGAATTCCTGAGTCAAGTTCATCTGAAGAGAGAGAAAACCTGATGATATGACTGG 969  
|||||  
QY 922 TTTGCTGTTAATCTCCAGATCACAACTGTACAGTGTACACAGTGTCTTATTTAGT 981  
|||||  
Db 970 TTTGCTGTTAATCTCCAGATCACAACTGTACAGTGTACACAGTGTCTTATTTAGT 1029  
|||||  
QY 982 GGAGCATTTATGTTAGAAATTCGAGCCAAAGTGGGAATGTACACAGTGTCTTATTTAGT 1041  
|||||  
Db 1030 GGAGCATTTATGTTAGAAATTCGAGCCAAAGTGGGAATGTACACAGTGTCTTATTTAGT 1089  
|||||  
QY 1042 AAGGCTGTAATGATAAAAAAGAACTGTCAACATTTACACAGTGTACATACAAATGCTGAG 1101  
|||||  
Db 1090 AAGGCTGTAATGATAAAAAAGAACTGTCAACATTTACACAGTGTACATACAAATGCTGAG 1149  
|||||  
QY 1102 ACAAATTTATACCTGGCAGAAAACCTACTGTTTGTGATTCATTTCCAAAGCTTATTCAATT 1161  
|||||  
Db 1150 ACAAATTTATACCTGGCAGAAAACCTACTGTTTGTGATTCATTTCCAAAGCTTATTCAATT 1209  
|||||





Db 2349 ACATTTTCAGACTCAATATAGACAGCTGTCTCATGTCTAAGACTGAGCAGAACTGAAA 2408  
QY 2362 AATTACTATTGGATATCACTCTTCTTTATATTTGTCATTGTGCACAACTTAAATAT 2421  
Db 2409 AATTACTATTGGATATCACTCTTCTTTATATTTGTCATTGTGCACAACTTAAATAT 2468  
QY 2422 ACTACCAAGTACAGAAATGTGGAAAAA 2453  
Db 2469 ACTACCAAGTACAAAAA 2500

RESULT 4  
PCT-US95-05008-3  
; Sequence 3, Application PC/TUS9505008  
; GENERAL INFORMATION:  
; APPLICANT: Sugen, Inc.  
; APPLICANT: 515 Galveston Drive  
; APPLICANT: Redwood City, California 94063-4720  
; APPLICANT: United States of America  
; APPLICANT: Wissenshaften E.V.  
; APPLICANT: Hofgarten Str. 2  
; APPLICANT: Munchen 80539  
; APPLICANT: Germany  
; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
; TITLE OF INVENTION: Kinases  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05008  
; FILING DATE: 24-APR-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/232,545  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212)869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2500 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA  
PCT-US95-05008-3

Query Match 97.6%; Score 2397.6; DB 5; Length 2500;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2417; Conservative 0; Mismatches 14; Indels 1; Gaps 1;  
QY 22 ACGGATGATATATGGATACAAATCTATTCTAGAAAGTCTTCTCAAAAGATCACAG 81  
Db 70 AAGSATGATATATGGATACAAATCTATTCTAGAAAGTCTTCTCAAAAGATCACAG 129  
QY 82 CAAAAGAGAAATGTCACCAAAATATACAAAGAACGGCTTTTGTGTTGACCAAAACA 141

Db 130 CAAAAGAGAAATGTCACCAAAATATACAAAGAACGGCTTTTGTGTTGACCAAAACA 189  
QY 142 AACCTTTCTCTACTATGATATGACAAATGAAAAGGGCAGCAGAAAAGAGTCCATTGAA 201  
Db 190 AACCTTTCTCTACTATGATATGACAAATGAAAAGGGCAGCAGAAAAGAGTCCATTGAA 249  
QY 202 ATTAAGAAAATCAGATGTGTGGAGAACTAAATCTCGAGGAGCAGACGCTGTAGAGAGA 261  
Db 250 ATTAAGAAAATCAGATGTGTGGAGAACTAAATCTCGAGGAGCAGACGCTGTAGAGAGA 309  
QY 262 CAGTACCCATTTCAGATTGTCTATAAAGATGGGCTTCTCTATGTCTATGCATCAATGAA 321  
Db 310 CAGTACCCATTTCAGATTGTCTATAAAGATGGGCTTCTCTATGTCTATGCATCAATGAA 369  
QY 322 GAGAGCCGAAGTCACTGTGTTGAAAGCATTACAAAAGAGATAAGGGGTAAACCCACCTG 381  
Db 370 GAGAGCCGAGTCACTGTGTTGAAAGCATTACAAAAGAGATAAGGGGTAAACCCACCTG 429  
QY 382 CTGCTCAAGTACCATTAGTGGTTCCTGCTGGACGGAAAGTTCCCTGCTGTGGCAGAGAGC 441  
Db 430 CTGCTCAAGTACCATTAGTGGTTCCTGCTGGACGGAAAGTTCCCTGCTGTGGCAGAGAGC 489  
QY 442 TGTAAAGCAGCCCGCAGGATGTACCTCTCTGGGAAGCATATGCTAAATCTGCATACTGCAGTC 501  
Db 490 TGTAAAGCAGCCCGCAGGATGTACCTCTCTGGGAAGCATATGCTAAATCTGCATACTGCAGTC 549  
QY 502 AATGAAGAGAAACACAGAGTTCCACCTTCCAGACAGAGTGTCTGAAGATACCTCGGGCA 561  
Db 550 AATGAAGAGAAACACAGAGTTCCACCTTCCAGACAGAGTGTCTGAAGATACCTCGGGCA 609  
QY 562 GTTCTCTCTCAAAATGGATGCACCATCTTCAAGTACCACCTCTCAAGTACCACCTCTGCAAC 621  
Db 610 GTTCTCTCTCAAAATGGATGCACCATCTTCAAGTACCACCTCTCAAGTACCACCTCTGCAAC 669  
QY 622 GAATCAAGAAATATGCTCCAGCCACCATCTTCAAGTACCACCTCTGCAAGTACCACCTCTGCAAG 681  
Db 670 GAATCAAGAAATATGCTCCAGCCACCATCTTCAAGTACCACCTCTGCAAGTACCACCTCTGCAAG 729  
QY 682 GACAGCACTCAAGAAATCTATGGCTCCAGCCACCATCTTCAAGTACCACCTCTGCAAGTACCACCT 741  
Db 730 GACAGCACTCAAGAAATCTATGGCTCCAGCCACCATCTTCAAGTACCACCTCTGCAAGTACCACCT 789  
QY 742 AGGGAAGACTTCCCTGACTGGTGGCAAGTAGAAAACCTGAAAAGTAGCAGCAGCAGTGA 801  
Db 790 AGGGAAGACTTCCCTGACTGGTGGCAAGTAGAAAACCTGAAAAGTAGCAGCAGCAGTGA 849  
QY 802 GATGTTGCAAGCAGTAACCAAAAAGAAAGAAATGTGAATCACACACCTCTCAAGATTTTCA 861  
Db 850 GATGTTGCAAGCAGTAACCAAAAAGAAAGAAATGTGAATCACACACCTCTCAAGATTTTCA 909  
QY 862 TGGGAATTCCTGACTCAAGTTTCACTGAAGAGAGAAAACCTGGATGATTGATTGACTGG 921  
Db 910 TGGGAATTCCTGACTCAAGTTTCACTGAAGAGAGAAAACCTGGATGATTGATTGACTGG 969  
QY 922 TTTGCTGGTAAACATCTCCAGATCACAATCTGAACAGTTTACTCAGACAAAAGGAAAAGAA 981  
Db 970 TTTGCTGGTAAACATCTCCAGATCACAATCTGAACAGTTTACTCAGACAAAAGGAAAAGAA 1029  
QY 982 GGAGCATTTATGGTTAGAAAATTCAGACCAAGTGGGAATGTACACAGTGTCTTTATTTAGT 1041  
Db 1030 GGAGCATTTATGGTTAGAAAATTCAGACCAAGTGGGAATGTACACAGTGTCTTTATTTAGT 1089  
QY 1042 AAGCTGTGAATGATATAAAGAAAGAACTGTCAAAATTTACCAGCTGCATACAATGCTGAG 1101  
Db 1090 AAGCTGTGAATGATATAAAGAAAGAACTGTCAAAATTTACCAGCTGCATACAATGCTGAG 1149  
QY 1102 AACAAATATACCTGGCAGAAAACCTACTGTTTGTATTCCTTCCAAAGCTTATTCATTAT 1161  
Db 1150 AACAAATATACCTGGCAGAAAACCTACTGTTTGTATTCCTTCCAAAGCTTATTCATTAT 1209  
QY 1162 CATCAACACAAATTCAGCAGGATGATCACAGGGTCCGCGACCTGTGTCAACAAGGCC 1221  
Db 1210 CATCAACACAAATTCAGCAGGATGATCACAGGGTCCGCGACCTGTGTCAACAAGGCC 1269



Qy	1222	AACAAAGTCCCGACTCTGTCTCCCTGGGAAATGGAATCTGGGAACCTGGAACCTGAAAAAGAGAAG	1281
Db	1270	AACAAGTCCCGACTCTGTCTCCCTGGGAAATGGAATCTGGGAACCTGGAACCTGAAAAAGAGAAG	1329
Qy	1282	ATTACCTTGTGAAGGAGCTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGCAAGTGG	1341
Db	1330	ATTACCTTGTGAAGGAGCTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGCAAGTGG	1389
Qy	1342	AAGGGCAGTATGATGTGTCTTAAAGTATGATCAAGGAGGECTCCATGTGCAGAAAGATGAA	1401
Db	1390	AAGGGCAGTATGATGTGTCTTAAAGTATGATCAAGGAGGECTCCATGTGCAGAAAGATGAA	1449
Qy	1402	TTCTTTCAGAGGCCACAGACTATGATGAACAACTCAGCCATCCCAAGCTGGTTAAATTCCTAT	1461
Db	1450	TTCTTTCAGAGGCCACAGACTATGATGAACAACTCAGCCATCCCAAGCTGGTTAAATTCCTAT	1509
Qy	1462	GGAGTGTGTCAAGAGGAATACCCCATATACATAGTCACTCAATATATAAGCAATGCGCTGC	1521
Db	1510	GGAGTGTGTCAAGAGGAATACCCCATATACATAGTCACTCAATATATAAGCAATGCGCTGC	1569
Qy	1522	TTGCTGAATTAACCTGAGGAGCTACGGGAAAAGGACTTGAACCTCCAGCTCTTAGAAATG	1581
Db	1570	TTGCTGAATTAACCTGAGGAGCTACGGGAAAAGGACTTGAACCTCCAGCTCTTAGAAATG	1629
Qy	1582	TGCTACGATCTGTGAAGGCATGGCCCTCTTGAGAGTCAACCAATTCATACACCCGGAC	1641
Db	1630	TGCTACGATCTGTGAAGGCATGGCCCTCTTGAGAGTCAACCAATTCATACACCCGGAC	1689
Qy	1642	TTGGCTGCTCGTAACCTGTTGGTGGACAGAGATCTCTGTGTAAGATATCTGACTTTTGA	1701
Db	1690	TTGGCTGCTCGTAACCTGTTGGTGGACAGAGATCTCTGTGTAAGATATCTGACTTTTGA	1749
Qy	1702	ATGACAGGATGTCTTGTGATGACCAGTATGTCACTTCAGTCGGACAAAGTTTCAGATC	1761
Db	1750	ATGACAGGATGTCTTGTGATGACCAGTATGTCACTTCAGTCGGACAAAGTTTCAGATC	1809
Qy	1762	AAGTGGTCAAGCTCCAGAGGTGTTTCATTCTCAATATACAGCAGCAAGTCAGACGATGG	1821
Db	1810	AAGTGGTCAAGCTCCAGAGGTGTTTCATTCTCAATATACAGCAGCAAGTCAGACGATGG	1869
Qy	1822	GCATTTGGGATCTGTATGTGGGAGGTGTTACAGCCTGGGGAAGCAGCCCTATGACTTGTAT	1881
Db	1870	GCATTTGGGATCTGTATGTGGGAGGTGTTACAGCCTGGGGAAGCAGCCCTATGACTTGTAT	1929
Qy	1882	GACAACCTCCAGAGTGTCTGTAAGGTCTCCAGGGCCACAGGCTTTACGGGCCCACTTG	1941
Db	1930	GACAACCTCCAGAGTGTCTGTAAGGTCTCCAGGGCCACAGGCTTTACGGGCCCACTTG	1989
Qy	1942	GCATCGGACACCATCTACCAAGATCATGTACAGCTGCTGGCAGCAGCTTCAGAAAAAGCGT	2001
Db	1990	GCATCGGACACCATCTACCAAGATCATGTACAGCTGCTGGCAGCAGCTTCAGAAAAAGCGT	2049
Qy	2002	CCACATTTTCAGCAACTCTGCTTCCATTTGAACCACTTCGGGAAAAAGACAAAGCATGA	2061
Db	2050	CCACATTTTCAGCAACTCTGCTTCCATTTGAACCACTTCGGGAAAAAGACAAAGCATGA	2109
Qy	2062	AGAAGAAATTAGAGTGTCTGATAAGAAATGAATATAGATGCTGGCCAGCATTTTCATTCAT	2121
Db	2110	AGAAGAAATTAGAGTGTCTGATAAGAAATGAATATAGATGCTGGCCAGCATTTTCATTCAT	2169
Qy	2122	TTTTAAGGAAATAGGAAGGCATAGTAATTTTAACTAGTTTTTTAATAGTGTCTCTGTAT	2181
Db	2170	TTTTAAGGAAATAGGAAGGCATAGTAATTTTAACTAGTTTTTTAATAGTGTCTCTGTAT	2229
Qy	2182	TGTCCTATTTTAAATTAACAAAGGCAGCAACAAAGATTCCTTGAATTTTATGATCA	2241
Db	2230	TGTCCTATTTTAAATTAACAAAGGCAGCAACAAAGATTCCTTGAATTTTATGATCA	2289
Qy	2242	AATTAGTAATTTTGTATCTGCTGCTCCCTGATATACACTTCCAGCCTATAGCAAGC	2301
Db	2290	AATTAGTAATTTGT - TTTATGCTGCCCTTGATATAACACTTTCCAGCCTATAGCAAGC	2348

## RESULT 5

US-08-391-615-1  
: Sequence 1, Application US/08391615  
: Patent No. 5550054  
: GENERAL INFORMATION:  
: APPLICANT: Witte, Owen  
: APPLICANT: Tsukada, Satoshi  
: APPLICANT: Saffran, Douglas  
: APPLICANT: Rawlings, David  
: TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE  
: TITLE OF INVENTION:  
: NUMBER OF SEQUENCES: 7  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT  
: STREET: 4 Embarcadero Center, Suite 3400  
: CITY: San Francisco  
: STATE: California  
: COUNTRY: USA  
: ZIP: 94111-4187  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/391,615  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/006,449  
: FILING DATE: 21-JAN-1993  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Rowland, Bertram I  
: REGISTRATION NUMBER: 20,015  
: REFERENCE/DOCKET NUMBER: A-57583-1/DIR UCLA 232-1  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 781-1989  
: TELEFAX: (415) 398-3249  
: TELEX: 910 277299 FHT UR  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 2505 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 137..2116  
US-08-391-615-1

	Query Match	20.1%	Score 493;	DB 1;	Length 2505;
	Best Local Similarity	63.9%	Pred. No. 1.4e-125;		
	Matches 745;	Conservative	0;	Mismatches 420;	Indels 0; Gaps 0;
QY	887	CTCAAGAAGAGAAAACCTGGATGATATATGACTGGTTTTGCTGCTAAACATCTCCAGATCAC	946		
Db	945	CTCAAGCTCGAGACTCCATAGAGATGTATGAGTGGTATTCCAAGACATGACTCGAAGTC	1004		
QY	947	AATCTGAACAGTTTACTCAGACAAAGGGAAAAGAGGAGCATTTATTTGGTTAGAAATTCGA	1006		



Db	1005	AGCTGAGCAACTGCTTAACAGACAGGGGAAAGAGAGGTTTCATTGTGTCAGAGACTCCA	1064
QY	1007	GCCAAAGTGGGAATGTACACAGTGTCCCTTATTTAGTAAGAGGCTGTGAATGATAAAAAAGGAA	1066
Db	1065	GCAAAGCTGGAAATAACACCGTGTCTGCTGTTTGGCTAAATCTACTGCGGAGCCTCAAGGG	1124
QY	1067	CTGTCAAAACATTACCAGGTGATACAAATGCTGTGAACAATAATATACCTGTGGCAGAAACT	1126
Db	1125	TGATCCGCCATTACGTTGTGTGTCTCCACGCCACAGAGCCAGTATTACCTGGCTGAGAAAC	1184
QY	1127	ACTGTTTTGATTCCATTCCAAAGCTTATTCATTATCATCAACAATTCACGAGGCATGA	1186
Db	1185	ACCTCTTCAGCACCATCCCTGAGCTCAATTAACATACCATCAACAACACTGTCAGGCGCTCA	1244
QY	1187	TCACACGCTCCGCGACCCCTGTGTCAACAAGGCCAACAAAGGTCGCCGACTCTGTGTGCC	1246
Db	1245	TATCCAGGCTGAATAATCCCTGTCTCTAAACAACAACAACAGCGCCCTTCTACTGCAAGCC	1304
QY	1247	TGGGAATGGAATCTTGGGAACCTGAAAGAGAGAGATTAACCTTGTGTAAGAGAGCTGGGAA	1306
Db	1305	TGGGCTATGGATCATGGGAAATGATCCAAAGGACCTCACCTCTTCAAGAGAGCTTGGGA	1364
QY	1307	GTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGGGAAGGGGCGAGTAGATGTTGCTGTTA	1366
Db	1365	CTGGCAAAATTCGGTGTCTGTAATATATGGGAAGTGGAGGGGCCAATATGATGTGGCCATCA	1424
QY	1367	AGATGATCAAGGAGGGCTCCATGTCAACAAGCTGAATTCCTTCAGGAGGCCACAGACTATGA	1426
Db	1425	AGATGATCAGAGAAGGTTCCATGTGGGAGGATGAATTCATATGAAGAAGCCAAAGTCATGA	1484
QY	1427	TGAACATCAACCCATCCCAAGCTGGTTAAATTCATATGGAGTGTGTTCAAAGGAATACCCCA	1486
Db	1485	TGAATCTTTCCTCATGAGAAGCTGGTGAGTTGATGCGCTCTGCACCAACAACAGCCCCA	1544
QY	1487	TATACATGTGACTGGAATATATAGCAATGCGCTTGTCTGTGAATTAACCTGAGGAGTCAGC	1546
Db	1545	TCCTCATCATCACCGAGTACATGSCATATGSGCTGCCCTCTTGAATCACTGAGGAGATGC	1604
QY	1547	GAAGAAGACTTGAACTTCCAGCTCTTACAAATGTGCTACGATGTCTGTGAAGGCATGG	1606
Db	1605	GGCACCGCTCCAGACACACAGCTGCTGAGATGTGCAAGAATGTCTGTGAAGCAATGG	1664
QY	1607	CCTTCTTTGAGAGTCAACAAATTCATACACGGGACTTGGCTGTGCTGTAACTGTGTGTGG	1666
Db	1665	AATACTTTGGAGTCGAAGCAGTTCCTTCACAGAGACCTGGCAGCTCGAAACTGTTGGTAA	1724
QY	1667	ACAGAGATCTCTGTGTGAAAGTATCTGACTTTTGGAAATGACAAGGTATGTTCTTGATGACC	1726
Db	1725	ACGATCAAGGAGTTGTGAAAGTATCTGACTTTGGCCCTGTCTAGTATGTCTCTGATGATG	1784
QY	1727	AGTATGTCACTTCAGTCGGGAACAAGTTTCCAGTCAAGTGGTCAAGTCCAGAGAGGTGTTTC	1786
Db	1785	AGTACACCACTCTGTAGGCTCCAAAGTTTCCAGTTCGGTGGTCTCCACCAAGAAGTGTTA	1844
QY	1787	ATTACTTCAAATACAGCAGCAAGTCAACAGCTATGGGCATTTGGGATCCCTGATGTGGGAGG	1846
Db	1845	TGTATAGCAAGTTCAGCAGCAAAATCTGCATCTGGGCTTTTGGGTTTTAATGTGGGAGA	1904
QY	1847	TGTTTCAGCTGGGAAGCAGCCCTATGACTTGTATGACAACTCCCAAGTGGTTCCTGAAGG	1906
Db	1905	TCCTACTCCCTTGGGAAGATCCGCTATGAGAGATTTTACTAACAGTGAACAGCAGCAAGACA	1964
QY	1907	TCCTCCAGGGCCACAGGCTTTACGGCCCACTGGCATCGGACATCGGACACCATCTACCAGATCA	1966
Db	1965	TTGCTCAAGGCTTACGTCTCTACAGGGCTCATCTGGCATCAGAGAGGTTATATACCATCA	2024
QY	1967	TGTACAGCTGCTGGCAGAGCTTCCAGAAAGGCTCCACATTTCAACAACTCCTGTCTT	2026
Db	2025	TGTACAGCTGCTGGCAGCAGAAAGCAGATGAACTCTAGTTTCAAAATTTCTCTTGAGTA	2084
QY	2027	CCATTGAACCACTTCGGGAAAAAGA	2051

DB 2085 ACATCTAGATGTGATGGAAGA 2109

RESULT 6  
US-09-142-529-2  
: Sequence 2, Application US/09142529  
: Patent No. 6225459  
: GENERAL INFORMATION:  
: APPLICANT: Mano, Hiroyuki  
: APPLICANT: Sakata, Tsuneaki  
: APPLICANT: Hasegawa, Mamoru  
: TITLE OF INVENTION: Promoter  
: FILE REFERENCE: 50026/011001  
: CURRENT APPLICATION NUMBER: US/09/142,529  
: CURRENT FILING DATE: 1999-08-12  
: EARLIER APPLICATION NUMBER: JP 8/54294  
: EARLIER FILING DATE: 1996-03-12  
: EARLIER APPLICATION NUMBER: PCT/JP97/00741  
: EARLIER FILING DATE: 1997-03-10  
: NUMBER OF SEQ ID NOS: 8  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 2  
: LENGTH: 2574  
: TYPE: DNA  
: ORGANISM: Mus musculus  
US-09-142-529-2

Query Match	17.9%;	Score 439.4;	DB 3;	Length 2574;
Best Local Similarity	61.5%;	Prod. No. 8.1e-111;		
Matches 723;	Conservative 0;	Mismatches 446;	Indels 6;	Gaps 1;
Qy	899	AAAACCTCGATGATATGACTGGCTTTGCTCGTAAACATCTCCAGATCACATCTCGACAGT	958	
Db	830	ACAACTTAGATCAATATGAGTGGTACTGCGAANAATACCAACGAAGCAAAAGCAGACAGC	889	
Qy	959	TACTCAGACAAAAGGGAAAGAGGACATTTATGTTAGAAATTCGAGCCAAAGTGGAA	1018	
Db	890	TCCTCAGACGGAAGATAAAGAGGTGGTTTTATGGTGACAGACTCCAGTCAACCCAGGCT	949	
Qy	1019	TGTACACAGTGTCTTATTTAGTAGAGCTGTGAATGATAAAAAGGAACGTGTCAACATTT	1078	
Db	950	TGTACACTGTCTCCCTTTACACAAAGTTTGGGGGAGAAAGGCTCATCAGGTTTTCAGGCAT	1009	
Qy	1079	ACCACGTGCA-----TACAAATGCTCAGAGAACAAATATACCTGGGAGAAACACTACTGTT	1132	
Db	1010	ATCACATAAAGGAAACAGCAACATCCCAAGAAGATATTACCTGGCAGAGAAGCATGCTTT	1069	
Qy	1133	TTGATTCATTCCAAAGCTTTATTCATTATCATCAACACAAATTCAGCAGGCGATGATCACAC	1192	
Db	1070	TCGGGTCCATTCTGTGAGATCATTTGAATATCACAAAGCACAAATGCGCGAGGCTTGTCACCA	1129	
Qy	1193	GGCTCCGCCACCTGTGTCAACAAAGGCCAACAGGTCCCGACTCTGTGTCTCCCTGGGAA	1252	
Db	1130	GGCTGCGGTACCGGTGAGTACAAGGGGAAGAACGCTCCCACTACTGCGGCCCTTCAGCT	1189	
Qy	1253	ATGGAATCTGGGAACCTAAAAGAGAGAGATTTACCTTGTGTGAAGGAGCTGGGAAGTGGCC	1312	
Db	1190	ATGATAGTGGGAGATTAAACCATCAGAGCTGACCTTTATGAGAGAGTTGGGAGCGGAC	1249	
Qy	1313	AGTTTGGAGTGGTCCAGCTGGGCAAGTGGGAAGGGGCGAGTATGATGTGTGTTTAAGATGA	1372	
Db	1250	TGTTTTGGAGTGGTGAAGCTTGGCAAGTGGCGGGCCAGTACAAGTGGCCATCAAAAGCTA	1309	
Qy	1373	TCAGGAGGGCTCCATGTCAGAGATGAATTCCTTCAGGAGGCCAGACTATCATGAAAC	1432	
Db	1310	TCGGGGAAGCGGCATGTGTGAGAGAGATTTTCATAGGAAGCTAAAGTCATGATGAAGC	1369	
Qy	1433	TCAGCCATCCAAAGCTGGTTAAATTCCTATGGAGTGTGTTCAAAGGAATACCCCATATACA	1492	
Db	1370	TGACACACCCCAAGCTGTACAGCTCTATGGTGTATGCACCCAGCAGAGCCCATCTACA	1429	
Qy	1493	TAGTGACTGAATATATAAAGCAATGGCTTGTGTAATTATTCAGGAGTCAACGGAAAG	1552	

Db 1430 TCGTTACCGAGTTTCATGGAACGGGGCTGCTTCTGAAATTCCTCCGGCAGACAGAGGCC 1489  
QY 1553 GACTTGAACCTTCCAGAGCTTTAGAAAATGCTGACGATGCTGTGGAAGGATGGCCCTTCT 1612  
Db 1490 ATTTCAGCAGAGACATGCTGCTAAAGCATGTCTCAAGATGCTGTGAAGGATGGAGTACC 1549  
QY 1613 TGGAGAGTCACCAATTCATACACGGGAGCTGGCTGCTGTAAGTCTGCTGGTGGACAGAG 1672  
Db 1550 TGGAGAGAACTTCTTCATCCACAGAGACTGGCTGCCAGGAATTTCTGATGAATGAAG 1609  
QY 1673 ATCTGTGTGAAAGTATCTGACTTTGGAATGACAAAGTATGTTCTTGTATGACCACTATG 1732  
Db 1610 CAGGAGTTGCTCAAGTATCTGATTTTGAATGGCAATGGCCAGGTACGTTCTGGATGATCAGTACA 1669  
QY 1733 TCAGTTCAGTCGGAACAAGTTTCCAGTCAAGTGGTCACTCCAGCTCCAGAGGTGTTTCATTACT 1792  
Db 1670 CAAGTTCTTCTTGGCCCAAGTTTCCCTGTGAAGTGGTGTCTCCCAAGAGTGTAAATTACA 1729  
QY 1793 TCAAAATACAGCAGCAAGTATGGCAATTTGGGATCTCTGATGTGGGAGGTCTTCA 1852  
Db 1730 GCGGCTTTAGCAGCAAGTATGAGTCTGCTGCTGTTGGTGTGCTTAATGTGGGAATATCA 1789  
QY 1853 GCGTGGGAAGCAGCGCTATGACTGTATGACAACTCCAGCTGCTGTTGAAAGTCTCC 1912  
Db 1790 CAGAGGGCAGGATGCGCTTTCAGAAAGACACCAATTAACGAAGTGGTAACTGGTGACTC 1849  
QY 1913 AGGCGCACAGGCTTTACCGGCCCCACCTGGCATCGGACCACTTACAGATCATGTACA 1972  
Db 1850 GTGGCCACCGCTTCCACCGGCCAAGCTGGCTTCCAAATATTTGTTATGAGTGTGCTGA 1909  
QY 1973 GCTGCTGGCAGCAGCTTCCAGAAAAGCGTCCACACATTTTCAGCAACTCTCTTCCATG 2032  
Db 1910 GATGCTGCGAAGAGAGACAGAGGAGGAGGCTTCTTGAAGACTTCTGCGTACGATAG 1969  
QY 2033 AACCACTTGGGAAAAGACAGCAATTTGAAGA 2067  
Db 1970 ATGAACCTAGTTGAATGTGAAGAACTTTTGGGAAGA 2004

## RESULT 7

US-10-045-428A-2  
; Sequence 2, Application US/10045428A  
; Patent No. 6500943  
; GENERAL INFORMATION:  
; APPLICANT: Mano, Hiroyuki  
; APPLICANT: Sakata, Tsuneaki  
; APPLICANT: Hasegawa, Mamoru  
; APPLICANT: Tabata, Toshiaki  
; TITLE OF INVENTION: Promoter  
; FILE REFERENCE: 50026/011003  
; CURRENT APPLICATION NUMBER: US/10/045,428A  
; CURRENT FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: 09/735,103  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 09/142,529  
; PRIOR FILING DATE: 1999-08-12  
; PRIOR APPLICATION NUMBER: PCT/JP97/00741  
; PRIOR FILING DATE: 1997-03-10  
; PRIOR APPLICATION NUMBER: JP 8/54294  
; PRIOR FILING DATE: 1996-03-12  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 2574  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-045-428A-2

Query Match 17.9%; Score 439.4; DB 4; Length 2574;  
Best Local Similarity 61.5%; Pred. No. 8.1e-111;  
Matches 723; Conservative 0; Mismatches 446; Indels 6; Gaps 1;  
QY 899 AAAACCTGGATATGACTGGTTGCTGTTAATCTCCAGATCACAATCTGAACAGT 958

Db 830 ACACCTTAGATCAATATGAGTGGTACTGCAAAATACCAACAGAAAGCAACAGACAGC 889  
QY 959 TACTCAGACAAAAGGAAAAGAGGAGCATTTATGTTAGAAAATTCAGGACCAAGTGGGAA 1018  
Db 890 TCCTCAGAACGGAAGATAAAGAGGTGTTTATGGTGAAGAGACTCCAGTCAACAGGCT 949  
QY 1019 TGTACACAGTGTCTCTTATTTAGTAAGGCTGTGAATGATAAAAAGGAAGACTGTCAACATTT 1078  
Db 950 TGTACACTGTCTCTCTTACACAAAGTTTGGGGAGAGGCTCATCAGGTTTCAGGCATT 1009  
QY 1079 ACCACCTGCA-----TACAAATGCTCAGAACAAATTTATACCTGGCAGAAAACACTCTGT 1132  
Db 1010 ATCACAATAAGGAACACAGCAACATCCCAAGAAAGATATTACCTGGCAGAGAGCATGCTT 1069  
QY 1133 TTGATTCCATTCCAAAGCTTTATTCATTATCATCAACACAAATTCAGCAGGATGATCAAC 1192  
Db 1070 TCGGCTCCATCTCTGAGATCATTTGAATATCACAAGCAATTCGCGCAGGGCTTGTCAACA 1129  
QY 1193 GGTCTCCGACCCCTGTCTCAAGAAAGGCCAACAAGTCCCGACTCTGTGTCCCTGGGAA 1252  
Db 1130 GGTCTCGGTACCCGGTCAAGAAAGGGAAGAACGCTCCCACTACTGCGGCCCTTCAGCT 1189  
QY 1253 ATGGAATCTGGGAACCTGAAAGAGAGAGATTTACCTTTGTTGAAGGAGCTGGGAAAGTGCC 1312  
Db 1190 ATGATAAGTGGGAGATTAACCCATCAGAGCTGACCTTTATGAGAGAGAGTGGGGAGCGGAC 1249  
QY 1313 AGTTTGGAGTGTCCAGCTGGGCAAGTGGGAAGGGCAGATGATGTTGCTGTGTTAAGATGA 1372  
Db 1250 TGTGTTGGAGTGTGGAGCTTGGCAAGTGGCGGCGCCAGTACAAAGTGGCCATCAAGCTA 1309  
QY 1373 TCNAGGAGGGCTCCATGTCAGAAAGTGAATTTCTTTCAGGAGGCCACAGACTATGATGAAC 1432  
Db 1310 TCCGGAAGCGCCCATGTGTGAAGAGGATTTTCATAGAGGAAGCTTAAATCATGATGAAGC 1369  
QY 1433 TCAGCCATCCCAAGCTGGTTAAATTTCTATGAGTGTCTTCAAGGAATACCCCATATATACA 1492  
Db 1370 TGCACACCCCAAGCTGGTACAGCTCTATGGTGTATGCACCCAGCAGAGCCCATCTACA 1429  
QY 1493 TAGTGACTGAATATATAAGCAANTGGCTGTGTTGTTGAATTTACCTGAGGAGTACCGGAAAG 1552  
Db 1430 TCGTTACCGAGTTTCATGGAACGGGCTGCTCTTCTGAAATTTCTCCCGCAGAGACAAGGCC 1489  
QY 1553 GACTTGAACCTTCCAGCTCTTAGAAAATGCTTACGATGCTGTGTAAGGATGAGCCCTTCT 1612  
Db 1490 ATTTACAGCAGAGACATGCTGCTAAGCATGTCTCAAGATGCTGTGAAGGATGAGTACC 1549  
QY 1613 TGGAGAGTCACCAATTTACACACCGGAGCTTGGTGTCTGCTCACTGCTTGGTGACACAGAG 1672  
Db 1550 TGGAGAGAACTTCTTCATCCACAGAGACCTGGCTGCCAGAAATTTCTAGTGAATGAAG 1609  
QY 1673 ATCTCTGTGTAAGTATCTGACTTTTGAATGACAAGGTATGTTCTTGTGATGACAGATG 1732  
Db 1610 CAGGAGTTCTCAAAGTATCTGATTTTGAATGGCCAGGTACGTTCTGGATGATCAGTACA 1669  
QY 1733 TCAGTTCACTGGGAACAAGTTTCCAGTCAAGTGGTTCAGCTCCAGAGGTGTTTCATTACT 1792  
Db 1670 CAAGTTCTTCTTCCGCAAGTTCCTCTGTAAGTGGTGTCCCAAGAGTGTGTTAAATTACA 1729  
QY 1793 TCAAAATACAGCAGCAAGTTCAGAGTATGGCAATTTGGGATCTCTGATGCTGGGAGGTGTTCA 1852  
Db 1730 GCGGCTTTAGCAGCAAGTTCAGAGCTGCTGGTGGTTTGTGTGCTAATTTGGGAATATATCA 1789  
QY 1853 GCGTGGGAAGCAGCGCTTATGACTTTGATGACAACTCCCAAGGTGGTTCTGAAGGTCTCCC 1912  
Db 1790 CAGAGGCGAGGATGCGCTTTGAGAAGAACCAATTAACGAAGTGGTAACTGGTGACTC 1849  
QY 1913 AGGCGCACAGGCTTTACCGGCCCCACCTGGCATCGGACCACTTACAGATCATGATACA 1972  
Db 1850 GTGGCCACCGCTTCCACCGGCCAAGCTGGCTTCCAAATATTTGTTATGAGTGTGCTGA 1909  
QY 1973 GCTGCTGGCAGCAGCTTCCAGAAAAGCGTCCCACTTTCAGCAACTCTCTTCCATG 2032

Db 1910 GATGCTGGCAAGAGAGACAGAGGAGGAGGCTTCTCTTTGAAGACTTGCTGCTAGATAG 1969  
QY 2033 AACCACTTCGGGAAAGACAGACGATTCAGACAAG 2067  
Db 1970 ATGACTAGTGTGAATGTGAAGAACTTTTGAAGA 2004

RESULT 8  
US-08-306-691B-35  
; Sequence 35, Application US/08306691B  
; Patent No. 5734039  
; GENERAL INFORMATION:  
; APPLICANT: Calabretta, Bruno  
; APPLICANT: Skorski, Tomasz  
; TITLE OF INVENTION: ANTISENSE  
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Seidel, Gonda, Lavorgna & Monaco, P.C.  
; STREET: Two Penn Center, Suite 1800  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306.691B  
; FILING DATE: September 15, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 8321-8  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; TELEX: No. 5734039e  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3623 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-306-691B-35

Query Match 9.0%; Score 220.2; DB 1; Length 3623;  
Best Local Similarity 52.9%; Pred. No. 2e-50;  
Matches 521; Conservative 0; Mismatches 458; Indels 6; Gaps 2;

QY 1061 AAGCAAGTCTCAACATACCACTGCATACAAATGCTGGAACAATATACCTGGCAG 1120  
Db 650 AAGGAGGGTGTACATACAGGATCAACATGCTGTGATGGCAAGCTCTACGTCTCT 709  
QY 1121 AAAAAGTCTGTTTGTATCCATTCCAAAGCTTATTCATCATCAACACAATTCAGCAG 1180  
Db 710 CCGAGAGCCGCTTCAACACCTGCCGAGTTGGTTTCATCATCACTCAACGGTGGCGAG 769  
QY 1181 GCATGATCAACAGGCTCGGCACCCCTGTGTCAACAAGGCAAGGCTCCCGACTCTG 1240  
Db 770 GGCTCATCAACAGGCTCCATTCACGCCCCAAGGCAAGGCAAGGCTCTATGGTG 829  
QY 1241 TGTCCTCGGGAATGGATCTGGGAATCAAGAGAGAGATACCTTGTGGAAGGAGC 1300  
Db 830 TGTCCTCGGGAATGGATCTGGGAATCAAGGAGATGGAAGGAGATCAAGGAGCAAGC 889  
QY 1301 TGGGAAGTGGCCAGTCTTGGAGTGTCTCCAGGCTGGGCAAGTGGAGGGCA---GTATGATG 1357

## RESULT 9

US-09-220-132-77

; Sequence 77, Application US/09220132

; Patent No. 6506607

; GENERAL INFORMATION:

; APPLICANT: Shyian, Andrew W.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT

; FILE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CA

; FILE REFERENCE: 07334-074001

; CURRENT APPLICATION NUMBER: US/09/220,132

; CURRENT FILING DATE: 1998-12-23

; PRIOR APPLICATION NUMBER: US 60/079,303

; PRIOR FILING DATE: 1998-03-25

; PRIOR APPLICATION NUMBER: US 60/068,821

; PRIOR FILING DATE: 1997-12-24

; NUMBER OF SEQ ID NOS: 191

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 77

; LENGTH: 2647

; TYPE: DNA

; ORGANISM: Homo sapiens

Db 890 TGGCGGGGGCCAGTACGGGAGGTGTACGAGGGCGTGTGGAAGAAATACAGCCTGACGG 949  
QY 1358 TTGCTGTTAAGATGATCAAGGAGGCTCCATGCTCAGAAGATGAATCTTTCAGGAGGCC 1417  
Db 950 TGGCGGTGAAGACCTTGAAGGAGGACACCATGGAGTGGAGAGGTCTTGAAGAGAGCTG 1009  
QY 1418 AGACTATGATGAACTCAGCCATCCAAAGCTGGTTAAATTTCTATGAGTGTCTTCAAAGG 1477  
Db 1010 CAGTCATGAAAGAGATCAACACCCCTAACCTAGTCAGCTCCTTGGGGTCTGCACCCGG 1069  
QY 1478 AATACCCCATATACATAGTGAATATATAAGCAATGGCTTGTGCTGAATACCTGA 1537  
Db 1070 AGCCCCCTTCTATATCATCATGATGATGATGATGATGATGATGATGATGATGATG 1129  
QY 1538 G---GAGTCACGGAAGAGCTTGAACCTTCCAGCTCTTAGAAATGTCTAGATGTCT 1594  
Db 1130 GGGAGTGAACGGCGGAGGTGAACGGCGTGTGCTGTGTACATGGCCACTCAGATCT 1189  
QY 1595 GTGAAGCATGGCTTCTTGGAGAGCTCACCATTATACACCGGAGCTTGGCTGCTCGTA 1654  
Db 1190 GTCAGCCATGAGTACCTAGAGAAGAAACTTATCCACAGATCTTGTCTGCCCGAA 1249  
QY 1655 ACTGCTTGGTGACAGAGATCTCTGTGAAAGTATCTGACTTTGGAATGACAAGGTATG 1714  
Db 1250 ACTGCTGTTAGGGGAGAACCACTTGTGAGGTAGTGTGATTTGGCCTGAGCAGTTGA 1309  
QY 1715 TTTTGTGACAGATATGTCAGTTTCAGTCGGGAAGAAAGTTTCCAGTCAAGTGTGAGTCT 1774  
Db 1310 TGACAGGGGACACCTTACACAGCCCATGCTGGAGCAAGTTCCCATCAATGGACTGCAC 1369  
QY 1775 CAGAGGTCTTTCATTACTTCAATACAGCAGCAAGTCAAGCTATGGGCTTGGCATCC 1834  
Db 1370 CCGAGAGCTGGCTTACAAACAAGTTCTCCATCAAGTCCAGCTCTGGGCTTGGAGTAT 1429  
QY 1835 TGATGTGGAGGTGTTCAGCCTGGGAAGCAGCCCTATGACTTGTATGACAACCTCCAGG 1894  
Db 1430 TGTCTTGGGAAATTCCTACCTATGTCATGTCCTTACCCGGGAATTACCGCTTCCAGG 1489  
QY 1895 TGGTCTCAAGGTCTCCAGGGCCAGGCTTACGGCCCACTGGGCTCGGACACCA 1954  
Db 1490 TGTATGAGTCTGTAGAGAAGGACTACCGCATGAAGCCCAAGGCTGCCCAAGAGG 1549  
QY 1955 TCTACAGATCATGTACAGCTGCTGGCAGAGCTTCCAGAAAAGCTCCACATTTTCAGC 2014  
Db 1550 TCTATGAACTATGCGAGCATGTTGGCAGTGAATCCCTCTGACCGCCCTCTCTTGTCTG 1609  
QY 2015 AACTCTCTCTTCCATTGAACCACT 2039  
Db 1610 AATCCCAAGCCCTTGAACAAT 1634



1520 AAGCGAGATCATGAAGAAGCTGAAGCAGCAGCAAGCTGGTCCAGCTCTATGCACTGCTGT 1579  
1472 CAAGAGATACCCATATACATAGTACTGAATATATATATATATATATATATATATATATAT 1531  
1580 CTGAGG---AGCCCATCATCATCTGCTACCGAGTATATGAACAAGAGAGTTTACTGGATT 1636  
1532 ACCTGA---GGAGTCAGGAGAAAGACTTGAACCTTCCAGCTCTTGAAGTATGCTACG 1588  
1637 TCTTAAGATGAGAGAGAGAGCTCTGAATATACCAATCTTGTGGATGCGACAC 1696  
1589 ATGCTGTGAAGCATGCGCTTCTTGGAGAGTACCAATTCATACACCGGAGCTTGGCTG 1648  
1697 AGGTGGCTGAGGAATGGCTTACATCAGCGCATGAATATATCATAGAGATCTGCGAT 1756  
1649 CTCGTRACTCTTGGTGGCAGAGATCTCTGTGTGAAGTATCTGACTTTGGATGACAA 1708  
1757 CAGCAACATTTAGTGGGAATGGACTCATATGCAAGATTTGCTGACTTCGGATTTGGCCC 1816  
1709 GGTATGTTCTTGATGACAGTATGTCAGTTCACTGGGAACAAGTTTCCAGTCAAGTGGT 1768  
1817 GATTGATAGAAGACAATGAGTACACAGCAAGAGCTGCAAGTTCCCATCAAGTGA 1876  
1769 CAGCTCAGAGGTGTTTCTTACTTCAATTAACATACAGCAGCAAGTACAGAGTATGGCATTTG 1828  
1877 CGCCCCCGAGGAGCGCTGACGGGAGTTTCAATCAAGTCTGACGTGTGGTCTTTTG 1936  
1829 GATCTGTGTTGGAGGTGTTCAAGCTGCGGAGAGAGCTTATGACTTGTATGACAACT 1888  
1937 GATCTTACTACAGAGCTGTTCAACAGAGAGAGTGCATACCCAGGAGATGACAACT 1996  
1889 CCAGGTGTTTGAAGTCTCCAGGCGCACAGGCTTTACCGGCCCCAGCTTGGCATCGG 1948  
1997 GGGAGGTGTTGGAGCGAGTGGAGCGAGCTACAGATGCGCTGCGCGAGGACTGCCCA 2056  
1949 ACACCTCTACAGATCATGATGATGCTGCGGACAGGCTTCCAGAGAAAGCTGCCACAT 2008  
2057 TCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2116  
2009 TTGAGCAACTCTGCTTCTTCCATGAA 2034  
2117 TTGAGTACTTGCAGAGCTTCTCTGGA 2142

## RESULT 11

US-08-391-615-7  
; Sequence 7, Application US/08391615  
; Patent No. 5550054  
; GENERAL INFORMATION:  
; APPLICANT: Witte, Owen  
; APPLICANT: Tsukada, Satoshi  
; APPLICANT: Saffran, Douglas  
; APPLICANT: Rawlings, David  
; TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/391,615  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/006,449

; FILING DATE: 21-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I  
; REGISTRATION NUMBER: 20,015  
; REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299 FHT UR  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1418 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-08-391-615-7

Query Match 8.5%; Score 207.6; DB 1; Length 1418;  
Best Local Similarity 62.5%; Pred. No. 3.6e-47;  
Matches 324; Conservative 0; Mismatches 194; Indels 0; Gaps 0;  
Qy 887 CTGAAGAGAGGAAACCTGGATGATTTATGACTGGTTTGGTGAACATCTCCAGATCAC 946  
Db 901 CTGAAGCAGAGAGACTCCATAGAAAATGTATGATGGTATTTCCAAACACATGACTCGGAGTC 960  
Qy 947 AATCTGAACAGTTACTCAGACAAAAGGAAAGAGAGGAGCATTTATGTTAGAAAATTCGA 1006  
Db 961 AGGCTGAGCAACTGCTTAAGCAAGAGGGAAGAGAGGTTTCATTGTGACAGACTCA 1020  
Qy 1007 GCCAAGTGGGAATGTACACAGTCTCTTTATTTAGTAAAGCTGTGAATGATAAAAAGGAA 1066  
Db 1021 GCAAGCTGCAAAATATACACTGCTGTGTTTGTCTAAATCCACAGGGGACCTCAAGGGG 1080  
Qy 1067 CTGTCAAACTACCACTGTCATACAAATGCTGAGACAAATATATACCTGGCAGAAACT 1126  
Db 1081 TGATACGTCTATTTGTTGTTTCCACACCTCAGAGCCAGTATTTACCTGGCTGAGAAGC 1140  
Qy 1127 ACTGTTTGTATCCATCCAAAGCTTATTCATATCATCAACACAAATTCAGCAGCATGA 1186  
Db 1141 ACCTTTTCAGCACCATCCCTGAGCTCATTAATCCATCAGCAGCAACTCTGCGAGACTCA 1200  
Qy 1187 TCACACGGCTCCGCGACCTGTGTCAACAAAAGGCCAACAGTCCCGACTCTGTGTCC 1246  
Db 1201 TATCCAGGCTCAAAATATCCAGTGTCTCAACAAAACAGAAATCCACCTTCCACTGCAAGCC 1260  
Qy 1247 TGGGAATGGAATCTGGAACTGAAAGAGAGAGATTAACCTTGTGAAGAGCTGGGAA 1306  
Db 1261 TGGGATACGGATCATGGGAAATTTGATCCAAAAGGACCTGACCTTCTTTGAAGGAGCTGGGGA 1320  
Qy 1307 GTGGCCAGTTTGGAGTGTCCAGCTGGGCAAGTGGAGGGGAGTATGATGTGTGCTGTTA 1366  
Db 1321 CTGCAAAATTTGGGTAGTAGTGAAGTATGGGAAATGGAGAGGCGCAGTACGACGTGGCCATCA 1380  
Qy 1367 AGATGATCAAGAGGGCTCCATGTCAGAGAAGTGAATTC 1404  
Db 1381 AGATGATCAAGAGGGCTCCATGTCGTAAGATGAATTC 1418

## RESULT 12

US-07-820-011A-3  
; Sequence 3, Application US/07820011A  
; Patent No. 5336615  
; GENERAL INFORMATION:  
; APPLICANT: Bell, Leonard  
; APPLICANT: Madri, Joseph A.  
; APPLICANT: Warren, Stephen L.  
; APPLICANT: Luthringer, Daniel J.  
; TITLE OF INVENTION: Genetically Engineered  
; TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced  
; TITLE OF INVENTION: Migration  
; TITLE OF INVENTION: and Plasminogen Activator Activity  
; NUMBER OF SEQUENCES: 4







```

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb storage
; OPERATING SYSTEM: IBM PC XT
; SOFTWARE: Displaywrite 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,011A
; FILING DATE: 19920106
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: LB-101
; TELEPHONE: (203) 255 1400
; TELEFAX: (203) 255 1401
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1602 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gallus, gallus
; PUBLICATION INFORMATION:
; AUTHORS: Tanaka, Hidesaburo
; TITLE: Structure and Sequence of the
; TITLE: Cellular Gene Homologous to the RSV src
; TITLE: Gene and the Mechanism for Generating the
; TITLE: Transforming Virus
; JOURNAL: Cell
; VOLUME: 32
; PAGES: 881-890
; DATE: March, 1983
;
US-07-820-011A-1
Query Match 8.3%; Score 204; DB 1; Length 1602;
Best Local Similarity 55.8%; Pred. No. 3.7e-46;
Matches 450; Conservative 0; Mismatches 350; Indels 6; Gaps 3;

QY 1231 CCGGACTCTGTGCTGGGAAATGGAATCTGGGAACTGAAAGAGAGAGATTTACCTTG 1290
DB 748 CCCAGAGCCAGGACTCGCGAGGCGCTGGGAAATCCCGGGAGTCTGCGGCTG 807
QY 1291 TTGAAGAGCTGGGAGTGGCCAGTTTGGAGTGTCCAGCTGGGCAAGTGAAGGGCAG 1350
DB 808 GAGGTGAAGTGGGAGGCTGTTGGAGAGTCTGGATGGGAGGACCTGGAAGGCGACC 867
QY 1351 TATGATGTTCTGTTAAGATGATCAGGAGGCTCCATGTCCAGAGATGAATTTTCAG 1410
DB 868 ACCAGAGTGGCCATAAGATCTGAAGCCCGCAACATGTCCCGGAGGCTCTCTGCG 927
QY 1411 GAGGCCAGACTATGATGAAGTCAAGCTCCCAAGCTGTTAAATCTATGAGTGTGT 1470
DB 928 GAAGCCCAAGTGAAGAGCTCCGCGATGAGAGCTGTTACGTGTACGCAAGTGTG 987
QY 1471 TCAAGAAATACCCCATATACATAGTGAATATATATAGCAATGCTGCTGCTGAAT 1530
DB 988 TC---GGAAGAGCCCATCTACATCGTCACTAGTACATGATGACGAAGGAGGCTCTGAT 1044
QY 1531 TACCTGA--GGAGTACAGGAAAGGACTTGAACCTTCC--CAGCTCTTAGAAATGTGTAC 1587
;
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 760 Kb storage
; COMPUTER: DELL 486/50
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: Displaywrite 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00445
; FILING DATE: 19930105
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/820,011
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: ALX-101PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255 1400
; TELEFAX: (203) 254 1101
; INFORMATION FOR SEQ ID NO: 1:
;
1045 TTCTGAAGGAGAGATGGGCAAGTACCTGCGGCTGCCACACACTCTCTCATATGCTGCT 1104
1588 GATGCTGTGAAGGATGCGCTTCTTGGAGAGTACCAATTCATACACCGGAGCTTGGCT 1647
1105 CAGATTGATCGGCGATGGCTTATCTGGAGAGAGTGAACCTACGTCACCGAGACCTGCGG 1164
1648 GCTGTAAGTCTGTTGGTGAGAGATCTCTCTGTGAAAAGTATCTGACTTTTGAATGACA 1707
1165 GCGCCAAACATCTGTTGGTGGGAGAACCTGTGTCAAGGTGGCTGACTTTGGGCTGGCA 1224
1708 AGGTATGTTCTGTATGACCAAGTATGTCAGTTTCAGTCGGAACAAAGTTTCCAGTCAAGTGG 1767
1225 GCGCTCATCGAGGACACAGAGTACACAGCAGGCAAGGTGCCAAGTTCCCATCAAGTGG 1284
1768 TCAGCTCCAGAGGTGTTTTCATTACTTCAATACAGCAGCAAGTTCAGACGATATGGGCAATT 1827
1285 ACAGCCCCGAGGAGCGCTCTATGCGCGGTTTCAACATCAAGTCGATGCTGTTGCTCTTC 1344
1828 GGGATCCTGATGTGGGAGGTGTTACGCTGGGAAAGCAGCCCTATGACTTTGATGACAC 1887
1345 GGCATCCTGCTGACTGAGCTGACCAACAGGCGCGGTGCCATACCCAGGAGTGGTCAAC 1404
1888 TCCAGGTGTTCTGAAGTCTCCAGGCGCACAGGCTTTACGCGGCCACCTGGCATCG 1947
1405 AGGAGGTGCTGGACCAAGTGGAGAGGCTTACCGCATGCTGCGCGCCGAGTCCGCC 1464
1948 GACACCATCTACAGATCATGTACAGTGTGCTGGCAGAGCTTCCAGAAAGGCTGCCACA 2007
1465 GAGTCGCTGATGACCTCATGTGCCAGTGTGCGGAGGAGGACCTTGAGGAGGCGGCCACT 1524
2008 TTTTCAGCAACTCTGCTTCCATTGA 2033
1525 TTTGAGTACCTGCAGGCGCTTCTCTGGA 1550
;
RESULT 15
PCT-US93-00445-1
; Sequence 1, Application PC/TUS9300445
; GENERAL INFORMATION:
; APPLICANT: Bell, Leonard
; APPLICANT: Madri, Joseph A.
; APPLICANT: Warren, Stephen L.
; APPLICANT: Luthringer, Daniel J.
; TITLE OF INVENTION: Genetically Engineered
; TITLE OF INVENTION: Endothelial Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 760 Kb storage
; COMPUTER: DELL 486/50
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: Displaywrite 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00445
; FILING DATE: 19930105
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/820,011
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: ALX-101PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255 1400
; TELEFAX: (203) 254 1101
; INFORMATION FOR SEQ ID NO: 1:
;

```



## SEQUENCE CHARACTERISTICS:

LENGTH: 1602 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: CDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Gallus, gallus  
PUBLICATION INFORMATION:  
AUTHORS: Takeya, Tatsuo  
Hidesaburo  
TITLE: Structure and Sequence of the  
Cellular Gene Homologous to the RSV src  
TITLE: Gene and the Mechanism for Generating the  
Transforming Virus  
JOURNAL: Cell  
VOLUME: 32  
PAGES: 881-890  
DATE: March, 1983  
PCT-US93-00445-1

Query Match 8.3%; Score 204; DB 5; Length 1602;

Best Local Similarity 55.8%; Pred. No. 3.7e-46;

Matches 450; Conservative 0; Mismatches 350; Indels 6; Gaps 3;

QY	1231	CCGACTCTGTGCCCTGGGAATGGAATCTGGAACTGAAAGAGAGATTACCTTG	1290
DB	748	CCCCACCCAGGAGCTCGCCAGAGCGCTGGGAAATCCCGGGAGCTGCGGCTG	807
QY	1291	TTGAAGGAGCTGGAGTGGAGTGGTCCAGCTGGCAGTGGCAAGTGGAGGCGC	1350
DB	808	GAGGTGAAGCTGGGGAGGCTGCTTGGAGAGTCTGGATGGGGAGCTGGAAGCGC	867
QY	1351	TATGATGTTCTGTTAAGATGATCAAGAGGGGTCCATGTGAGAAGATGAATTTTC	1410
DB	868	ACCAGAGTGCCCATTAAGACTCTGAAGCCCGGCAACATGTCCCGGAGGCTTCTC	927
QY	1411	GAGGCCAGACTATGATGAAGTCAAGCTAGCCATCCCAAGCTGGTTAAATTTCT	1470
DB	928	GAAGCCCAAGTATGAAGAAGCTCCGGCATGAGAAGCTGTTACGCTGTACGAGT	987
QY	1471	TCAAAGGAATACCCATATACATGACCTGAATATATATAGCATGCTGCTTCTGA	1530
DB	988	TC---GGAAGAGCCCATCTACATCGTCACTGAGTACATGAGCAAGGGAGGCT	1044
QY	1531	TACCTGA--GGAGTACGGAAAAGGACTTGAACCTTCC--CAGCTCTTTAGAAAT	1587
DB	1045	TTCTGAAGGAGAGATGGCAAGTACTCGCGCTGCCACAGCTCGTGCATATGGCT	1104
QY	1588	GATGTCGTGAAGCATGGCTTCTTGGAGAGTCAACAAATCATACACGGGAGCTTG	1647
DB	1105	CAGATTGCATCCGGCATGGCTATGTGGAGAGGATGAACACTACGTGCACCGAG	1164
QY	1648	GCTCGTAACCTGTGGTGACAGATCTCTGTGAAAGTATCTGACTTTGGAAATGA	1707
DB	1165	CGGGCAACATCTGGTGGGGAGAACCTGGTGTGCAAGGTGGCTGACTTTGGGCT	1224
QY	1708	AGGTATGTTCTGATACCATATGTCAAGTTCAAGTGGGACAAAGTTTCCAGTCA	1767
DB	1225	CGCTCATCAGGACACAGAGTACACAGCAGCGCAAGGTGCCAAGTTTCCCATC	1284
QY	1768	TCAGCTCCAGAGGTGTTTACTTCAATATACAGCAGCAAGTCAAGCTATGGGCA	1827
DB	1285	ACAGCCCCCGAGGCGCCCTATATGGCGGGTTTACCATCAAGTCGGATGTCTGT	1344
QY	1828	GGATCCTGATGTTGGAGGTGTTTCAGCTGGGGAGCAGCCCTATGACTTGTATGA	1887
DB	1345	GGATCCTGCTGATGAGCTGACCAACCAAGGGCGGGTGGCCATACCCAGGGAT	1404
QY	1888	TCCAGGTGTTCTGAAGGTCTCCCGAGGCGCACAGGCTTTACCGGCCCCACCT	1947

Db	1405	AGGAGGCTGTGGACCAAGTGGAGGGGCTACCGCATGCCCTGCCCGCCGAGTGGCCC	1464
QY	1948	GACACCATCTACCAAGATCATGTGCTGGCAGCAGCTTCCAGAAAAGCGTCCCA	2007
Db	1465	GAGTGGCTGCATGACCTCATGTGCCATGTCTGGGGAGGGACCTTGAGAGCGG	1524
QY	2008	TTTCAGCAACTCCCTGTCTTCCATTGA	2033
Db	1525	TTTGAGTACCTGCAGGCGCTTCTCTGGA	1550

Search completed: August 20, 2003, 16:30:06  
Job time : 170 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: August 20, 2003, 15:10:22 ; Search time 556 Seconds

(without alignments)  
9930.738 Million cell updates/sec

Title: US-10-021-660-41  
Perfect score: 2456  
Sequence: 1 gcaagcagcagcaagctga.....aatgtggaataaaacacg 2456

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 1517243 seqs, 1124081882 residues  
Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2456	100.0	2456	12 US-10-021-660-41	Sequence 41, Appl
2	2456	100.0	2456	12 US-10-021-660-41	Sequence 41, Appl
3	2424.4	98.7	2415	11 US-10-021-660-41	Sequence 1, Appl
4	2404.2	97.9	2604	14 US-10-171-581-317	Sequence 11, Appl
5	2397.6	97.6	2500	9 US-09-977-269-3	Sequence 317, Appl
6	2397.6	97.6	2500	10 US-09-977-269-3	Sequence 3, Appl
7	2397.6	97.6	2500	11 US-09-977-269-3	Sequence 3, Appl
8	650.8	26.5	797	11 US-09-955-999-27	Sequence 27, Appl
9	494.6	20.1	2468	14 US-10-045-202-3	Sequence 3, Appl
10	485	19.6	2560	14 US-10-045-202-1	Sequence 1, Appl
11	480.2	19.6	2582	14 US-10-220-801-4	Sequence 4, Appl
12	439.4	17.9	2574	9 US-09-735-103-2	Sequence 2, Appl
13	439.4	17.9	2574	13 US-10-045-428A-2	Sequence 2, Appl
14	394.4	16.1	3663	10 US-09-919-172-84	Sequence 84, Appl
15	388	15.8	3593	14 US-10-220-801-6	Sequence 6, Appl
16	385.4	15.7	4221	14 US-10-220-801-8	Sequence 8, Appl

17	385.4	15.7	6383	10 US-09-954-531-405	Sequence 405, Appl
18	383.8	15.6	6381	12 US-10-007-926A-45	Sequence 45, Appl
19	263.4	10.7	431	14 US-10-106-698-1335	Sequence 1335, Ap
20	223.4	9.1	5527	10 US-09-880-107-3710	Sequence 3710, Ap
21	211.6	8.6	4414	12 US-10-101-510-512	Sequence 512, Appl
22	208	8.5	347	10 US-09-833-381-1818	Sequence 1818, Ap
23	208	8.5	347	10 US-09-833-381-1819	Sequence 1819, Ap
24	191.4	7.8	2354	10 US-09-967-768A-300	Sequence 300, Appl
25	189.2	7.7	2015	10 US-09-954-456-1983	Sequence 1983, Ap
26	189.2	7.7	2015	14 US-10-007-010-3	Sequence 3, Appl
27	181.2	7.4	2298	14 US-10-175-523-50	Sequence 50, Appl
28	178	7.2	1911	10 US-09-917-800A-1611	Sequence 1611, Ap
29	168	6.8	2770	9 US-09-977-269-5	Sequence 5, Appl
30	168	6.8	2770	10 US-09-977-269-5	Sequence 5, Appl
31	168	6.8	2770	11 US-09-977-269-5	Sequence 5, Appl
32	168	6.8	2863	10 US-09-954-456-1691	Sequence 1631, Ap
33	168	6.8	7607	10 US-09-982-610-19	Sequence 19, Appl
34	160.2	6.5	454	11 US-09-918-995-26678	Sequence 20678, A
35	159	6.5	462	9 US-09-864-765-558	Sequence 558, Ap
36	158.8	6.5	347	10 US-09-796-692-8836	Sequence 8836, Ap
37	158.8	6.5	347	14 US-10-040-662-8836	Sequence 8836, Ap
38	157.6	6.4	774	11 US-09-956-622A-6	Sequence 6, Appl
39	153	6.2	4517	12 US-10-007-926A-304	Sequence 304, Appl
40	153	6.2	4517	15 US-10-298-377A-3	Sequence 3, Appl
41	148.2	6.0	452	9 US-09-864-761-11293	Sequence 11293, A
42	142.6	5.8	2674	13 US-10-003-295-1	Sequence 1, Appl
43	138.4	5.6	2451	10 US-09-771-161A-4	Sequence 4, Appl
44	136.2	5.5	1574	9 US-09-870-962-12	Sequence 12, Appl
45	134	5.5	159	11 US-09-158-722-15	Sequence 15, Appl

## ALIGNMENTS

### RESULT 1

US-10-021-660-41  
Sequence 41, Application US/10021660  
Publication No. US20030152926A1  
GENERAL INFORMATION:  
APPLICANT: Murray, Richard  
APPLICANT: Glynnne, Richard  
APPLICANT: Watson, Susan R.  
APPLICANT: EOS Biotechnology, Inc.  
TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis  
TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis  
FILE REFERENCE: 018501-000710US  
CURRENT APPLICATION NUMBER: US/10/021,660  
CURRENT FILING DATE: 2001-12-06  
PRIOR APPLICATION NUMBER: US/09/784,356  
PRIOR FILING DATE: 2001-02-14  
PRIOR APPLICATION NUMBER: US/09/637,977  
PRIOR FILING DATE: 2000-08-11  
NUMBER OF SEQ ID NOS: 135  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 41  
LENGTH: 2456  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-021-660-41

Query Match	100.0%;	Score 2456;	DB 12;	Length 2456;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches-2456;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	GCACGACGGACACAGCTGAGACGGATGATATATATCGATACAAAAATCTATTCTAGAAGAA	60	
Db	1	GCACGACGGACACAGCTGAGACGGATGATATATCGATACAAAAATCTATTCTAGAAGAA	60	
Oy	61	CTTCTTCTCAAAAGATTCACAGCAAGAAGAAAAATGTCACCAAAATTAATTACAAGAACGG	120	
Db	61	CTTCTTCTCAAAAGATTCACAGCAAGAAGAAAAATGTCACCAAAATTAATTACAAGAACGG	120	

Qy	121	CTTTTGTGTTTGACCAAAACAAACCTTTTCCTACTATGATGATGACAAAATGAAAGGGGG	180
Db	121	CTTTTGTGTTTGACCAAAACAAACCTTTTCCTACTATGATGATGACAAAATGAAAGGGGG	180
Qy	181	AGCAGAAAAGGATCCATTGAAATTAAGAAAATCAGATGTGTGGAGAAAGTAAATCTCGAG	240
Db	181	AGCAGAAAAGGATCCATTGAAATTTAGAAAATCAGATGTGTGGAGAAAGTAAATCTCGAG	240
Qy	241	GAGCAGACGCTGTAGAGAGACAGTACCCCATTTTCAGATTGTCTATAAAGATGGGCTTCTC	300
Db	241	GAGCAGACGCTGTAGAGAGACAGTACCCCATTTTCAGATTGTCTATAAAGATGGGCTTCTC	300
Qy	301	TATGCTATGCAATCAATGAAGAGAGCGGAAGTCAGTGGTTGGAAGCATTTACAAAAGAG	360
Db	301	TATGCTATGCAATCAAAATGAAGAGAGCGGAAGTCAGTGGTTGGAAGCATTTACAAAAGAG	360
Qy	361	ATAAGGGTAAACCCCACTGCTGGTCAAGTACCATAGTGGTTCCTTCGTGGACGGGAAG	420
Db	361	ATAAGGGTAAACCCCACTGCTGGTCAAGTACCATAGTGGTTCCTTCGTGGACGGGAAG	420
Qy	421	TTCTGTGTTGCCACAGAGCTGTAAAGCAGCCCCCAGGATGTACCCCTCTGGGAAGCATAT	480
Db	421	TTCTGTGTTGCCACAGAGCTGTAAAGCAGCCCCCAGGATGTACCCCTCTGGGAAGCATAT	480
Qy	481	GCTAATCTGCATACTGCAGTCAATGAAGAGAAACACAGAGTTCCCACTTCCCAACAGAGA	540
Db	481	GCTAATCTGCATACTGCAGTCAATGAAGAGAAACACAGAGTTCCCACTTCCCAACAGAGA	540
Qy	541	GTGCTGAAGATACCTTCGGGCAAGTTCCTGTTCTCAAAATGGATGCACCATCTTCAAGTACC	600
Db	541	GTGCTGAAGATACCTTCGGGCAAGTTCCTGTTCTCAAAATGGATGCACCATCTTCAAGTACC	600
Qy	601	ACTTAGCCCAATATGACACGAATCAAGAGAAAACCTATGGCTCCCAAGCACCATTCTTCA	660
Db	601	ACTTAGCCCAATATGACACGAATCAAGAGAAAACCTATGGCTCCCAAGCACCATTCTTCA	660
Qy	661	AGTACCAGTCTAGGGCAATATGACAGCAACTCAAGAAAATCTATGCTCCCAAGCACC	720
Db	661	AGTACCAGTCTAGGGCAATATGACAGCAACTCAAGAAAATCTATGCTCCCAAGCACC	720
Qy	721	TTCAACATGCAGTATATCCAAAGGAGACTTCCCTGACTGGTGGCAAGTAAGAAACTG	780
Db	721	TTCAACATGCAGTATATCCAAAGGAGACTTCCCTGACTGGTGGCAAGTAAGAAACTG	780
Qy	781	AAAAGTAGCAGCAGCTGAAGATGTTGCAAGCAGTAAACCAAAAGAAAAGAAATGTCAAT	840
Db	781	AAAAGTAGCAGCAGCTGAAGATGTTGCAAGCAGTAAACCAAAAGAAAAGAAATGTCAAT	840
Qy	841	CACACACCTCAAAAGATTTGATGGGAATTCCTGAGTCAAGTTCATCTGAAGAGAGGAA	900
Db	841	CACACACCTCAAAAGATTTGATGGGAATTCCTGAGTCAAGTTCATCTGAAGAGAGGAA	900
Qy	901	AACCTGATGATATGACTGGTTGCTGGTAACATCTCCAGATCCCAATCTGACACGTTA	960
Db	901	AACCTGATGATATGACTGGTTGCTGGTAACATCTCCAGATCCCAATCTGACACGTTA	960
Qy	961	CTCAGACAAAAGGGAAGAGCAGCATTTATGGTTAGAAATTCGAGCCCAAGTGGGAATG	1020
Db	961	CTCAGACAAAAGGGAAGAGCAGCATTTATGGTTAGAAATTCGAGCCCAAGTGGGAATG	1020
Qy	1021	TACACAGTGCCTTATTTAGTAGGCTGTGAATGATAAAAAGGAACGTGCAACATTTAC	1080
Db	1021	TACACAGTGCCTTATTTAGTAGGCTGTGAATGATAAAAAGGAACGTGCAACATTTAC	1080
Qy	1081	CAGTGATACAAATGCTGAGACAAATTTATACCTGGCAGAAAACCTACTGTTTGTATTC	1140
Db	1081	CAGTGATACAAATGCTGAGACAAATTTATACCTGGCAGAAAACCTACTGTTTGTATTC	1140
Qy	1141	ATTCCAAAGCTTATTTCAATATCATCAACAAATTCACGACGCGATGATCACAGGCTCCG	1200
Db	1141	ATTCCAAAGCTTATTTCAATATCATCAACAAATTCACGACGCGATGATCACAGGCTCCG	1200
Qy	1201	CAGCCTGTGTCACAAAAGGCCAACAGAGTCCCGACCTGTGTCCTCCCTGGGAAATGGAATC	1260

1201	 CACCCCTGTGTTCAACAAGGCCAACAAAGTCCCGAGTCTGTGTGCCCTGGGAATTGGAAATC	1360
1261	TGGGAATCTGAAAAGAGAAGAGATTACCTTGTTGCAAGGAGCTGGGAAGTGGCCAGTTTGGGA	1320
1261	TGGGAATCTGAAAAGAGAAGAGATTACCTTGTTGAGGAGCTGGGAAGTGGCCAGTTTGGGA	1320
1321	GTGGTCCAGCTGGGCAAGTGGGAAGGGGCAGTATGATGTTGCTGCTTTAAGATGATCAAGGAG	1380
1321	GTGGTCCAGCTGGGCAAGTGGGAAGGGGCAGTATGATGTTGCTGTTAAGATGATCAAGGAG	1380
1381	GGCTCCATGTCAGAAGATGAATTTCTTTTCAGGAGGCCAGACTATGATGAAATCTACGCCAT	1440
1381	GGCTCCATGTCAGAAGATGAATTTCTTTTCAGGAGGCCAGACTATGATGAAATCTACGCCAT	1440
1441	CCCAAGCTGTTTAAATTTCTATGAGGTGTGTTCAAAAGGAATACCCCATATACATAGTGACT	1500
1441	CCCAAGCTGTTTAAATTTCTATGAGGTGTGTTCAAAAGGAATACCCCATATACATAGTGACT	1500
1501	GAATATATAGCAATGGCTGCTTTGCTGTAATTACCTGAGGAGTCAAGGAAAAGACTTTGAA	1560
1501	GAATATATAGCAATGGCTGCTTTGCTGTAATTACCTGAGGAGTCAAGGAAAAGACTTTGAA	1560
1561	CTTTCGCCAGCTTTAGAAATGTGCTACGATGTCTGTGAAGGCATGGCCCTTCTTTGGAGAGT	1620
1561	CTTTCGCCAGCTTTAGAAATGTGCTACGATGTCTGTGAAGGCATGGCCCTTCTTTGGAGAGT	1620
1621	CACCAATTCATACACGGGACCTTGGCTGCTGCTGTAACCTGCTGGTGGACAGAGATCTCTGT	1680
1621	CACCAATTCATACACGGGACCTTGGCTGCTGCTGTAACCTGCTGGTGGACAGAGATCTCTGT	1680
1681	GTGAAAGTATCTGACTTTTGGAAATGACAAGGATATGTTCTTGATCACCAGTATGTCAGTTCA	1740
1681	GTGAAAGTATCTGACTTTTGGAAATGACAAGGATATGTTCTTGATCACCAGTATGTCAGTTCA	1740
1741	GTGCGGAACAAGTTTCCAGTCAAGTGGTCAGCTCCAGAGGTGTTTCAATTCAAAATAC	1800
1741	GTGCGGAACAAGTTTCCAGTCAAGTGGTCAGCTCCAGAGGTGTTTCAATTCAAAATAC	1800
1801	AGCAGCAAGTCAACAGCTATGGGCAATTTGGGATCCCTGATGTGGGAGGTGTTCAAGCTGGGG	1860
1801	AGCAGCAAGTCAACAGCTATGGGCAATTTGGGATCCCTGATGTGGGAGGTGTTCAAGCTGGGG	1860
1861	AAGCAGCCCTATGACTTTGTATGACAACTCCCAGTGGTCTGTAAGGTCTCCCCAGGGCCAC	1920
1861	AAGCAGCCCTATGACTTTGTATGACAACTCCCAGTGGTCTGTAAGGTCTCCCCAGGGCCAC	1920
1921	AGGCTTTACGGGCCACCTGGCATCGGACACCATCTACCAGATCATGTACAGCTGCTGG	1980
1921	AGGCTTTACGGGCCACCTGGCATCGGACACCATCTACCAGATCATGTACAGCTGCTGG	1980
1981	CACGAGCTTCCAGAAAGCGTCCACATTTACAGCAACTCCTGTCTTCCATTTGAACCACTT	2040
1981	CACGAGCTTCCAGAAAGCGTCCACATTTACAGCAACTCCTGTCTTCCATTTGAACCACTT	2040
2041	CGGGAAGAACAGCAAGCATTTGAAGAAGAAATAGGAGTGTGTATGAAGATGAATATAGATG	2100
2041	CGGGAAGAACAGCAAGCATTTGAAGAAGAAATAGGAGTGTGTATGAAGATGAATATAGATG	2100
2101	CTGGCCAGCAATTTTTCATTTTCAATTTTAAAGAAAGTAGGAAGGCATTAAGTAATTTAGCTAGT	2160
2101	CTGGCCAGCAATTTTTCATTTTCAATTTTAAAGAAAGTAGGAAGGCATTAAGTAATTTAGCTAGT	2160
2161	TTTTTAATAGTGTCTCTGTATGTCCTATATTTAGAAATGCAACAGCGAGGAACAAAG	2220
2161	TTTTTAATAGTGTCTCTGTATGTCCTATATTTAGAAATGCAACAGCGAGGAACAAAG	2220
2221	ATTCCCTTGAAATTTTAGATCAAAATTAGTAATTTGTTTATGCTGCTCTGATATAACAC	2280
2221	ATTCCCTTGAAATTTTAGATCAAAATTAGTAATTTGTTTATGCTGCTCTGATATAACAC	2280
2281	TTTTCCAGCCTATAGCAGAAGCACATTTTCAGACTGCAATATAGAGACTGTGTTTCATGTGT	2340

```

Db      2281  TTCCAGCCCTATAGCAGAAGCACATTTTCAGACTGCAATATAGAGACTGTGTTCAATGTG 2340
Qy      2341  AAAGACTGAGCAGAACTGAAAAATTAATTGATATTGATATTGATCTTTTCTTTATATTGTC 2400
Db      2341  AAAGACTGAGCAGAACTGAAAAATTAATTGATATTGATCTTTTCTTTATATTGTC 2400
Qy      2401  ATTGTCACACAATTAATAATTACTACCAAGTACAGAAATGTGGAAAAAAAACCCG 2456
Db      2401  ATTGTCACACAATTAATAATTACTACCAAGTACAGAAATGTGGAAAAAAAACCCG 2456

RESULT 2
US-10-186-399-1
; Sequence 1, Application US/10186399
; Publication NO. US20020173481A1
; GENERAL INFORMATION:
; APPLICANT: Ekman, Niklas
; APPLICANT: Arighi, Elena
; APPLICANT: Vastrik, Imre
; APPLICANT: Tamagnone, Luca
; APPLICANT: Alitalo, Kari
; TITLE OF INVENTION: REGULATION OF VASCULAR ENDOTHELUM USING BMX TYROSINE
; TITLE OF INVENTION: KINASE
; FILE REFERENCE: 28113/31941A
; CURRENT APPLICATION NUMBER: US/10/186,399
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 08/320,432
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1:
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-186-399-1

```

481	Db	 GCTAACTGCACTACTGTCAGTCAATGAAAGAGAAAACACAGAGTTCCTCACCTTCCCGACAGAGA	540
541	Qy	 GTGCTGAAGATACCTTCGGGAGATTCCCTGTTCTCAAAATGGATGATGCACCATCTTCAAGTACC	600
541	Db	 GTGCTGAAGATACCTTCGGGAGATTCCCTGTTCTCAAAATGGATGATGCACCATCTTCAAGTACC	600
601	Qy	 ACTCTAGCCCAATATGACAAAGCAATCAAGAAAACCTATGGCTCCAGCCACCACCTCTTCA	660
601	Db	 ACTCTAGCCCAATATGACAAAGCAATCAAGAAAACCTATGGCTCCAGCCACCACCTCTTCA	660
661	Qy	 AGTACCAAGTCTAGGGCAATATGACAGCAACTCAAGAAAATCTATGGCTCCAGCCCAAAAC	720
661	Db	 AGTACCAAGTCTAGGGCAATATGACAGCAACTCAAGAAAATCTATGGCTCCAGCCCAAAAC	720
721	Qy	 TTCAACATGAGTATATTCCAAGGGAAGACTTCCCTGACTGGTGGCAAGTAAGAAAACCTG	780
721	Db	 TTCAACATGAGTATATTCCAAGGGAAGACTTCCCTGACTGGTGGCAAGTAAGAAAACCTG	780
781	Qy	 AAAAAGTACGACGACGATGAAATTTCAAGGCAAGTAAACCAAGCAAGTAAACCAAGCAAGTAAAC	840
781	Db	 AAAAAGTACGACGACGATGAAATTTCAAGGCAAGTAAACCAAGCAAGTAAACCAAGCAAGTAAAC	840
841	Qy	 CACACCACTTCAAGATTTTCATGGGAATTCCTCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	900
841	Db	 CACACCACTTCAAGATTTTCATGGGAATTCCTCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	900
901	Qy	 AACCTGGATGATTTATGACTGTTTGGTGAACATCTCCAGATCACAATCTGAACAGTTA	960
901	Db	 AACCTGGATGATTTATGACTGTTTGGTGAACATCTCCAGATCACAATCTGAACAGTTA	960
961	Qy	 CTCAGACAAAGGGAAGGAGAGCAATTTATGGTTAGAAATTCGAGCCAAAGTGGGAATG	1020
961	Db	 CTCAGACAAAGGGAAGGAGAGCAATTTATGGTTAGAAATTCGAGCCAAAGTGGGAATG	1020
1021	Qy	 TACACAGTGTCTTATTTAGTAAAGCTGTGAATGATAAAAAGCAACTGTCAAAACATTTAC	1080
1021	Db	 TACACAGTGTCTTATTTAGTAAAGCTGTGAATGATAAAAAGCAACTGTCAAAACATTTAC	1080
1081	Qy	 CACGTGCATACAAATGCTGAGCAAAATATACCTGGCAGAAAACTACTGTTTGTATTTCC	1140
1081	Db	 CACGTGCATACAAATGCTGAGCAAAATATACCTGGCAGAAAACTACTGTTTGTATTTCC	1140
1141	Qy	 ATTCCAAAGCTTATTCATATATCATCACACAATTCAGCAGGCATGATCACACGGCTCCGC	1200
1141	Db	 ATTCCAAAGCTTATTCATATATCATCACACAATTCAGCAGGCATGATCACACGGCTCCGC	1200
1201	Qy	 CACCTGTGTCAACAAAGGCCAACAGGTCCCCGACTCTGTGTCCCTGGGAATTTGGATTC	1260
1201	Db	 CACCTGTGTCAACAAAGGCCAACAGGTCCCCGACTCTGTGTCCCTGGGAATTTGGATTC	1260
1261	Qy	 TGGGAAGCTTAAAGAGAGAGATTTACCTGTTTGAAGGAGCTGGGAAGTGGCCAGCTTTTGA	1320
1261	Db	 TGGGAAGCTTAAAGAGAGAGATTTACCTGTTTGAAGGAGCTGGGAAGTGGCCAGCTTTTGA	1320
1321	Qy	 GTGCTCCAGCTGGGCAAGTGGAAAGGCGCATGATGATGCTGTTTGAAGATGATCAAGAG	1380
1321	Db	 GTGCTCCAGCTGGGCAAGTGGAAAGGCGCATGATGATGCTGTTTGAAGATGATCAAGAG	1380
1381	Qy	 GGCTCCATGTCAAGATGATTTCTTTCAGGAGGCCAGACTATGATGAAACTCAGCCAT	1440
1381	Db	 GGCTCCATGTCAAGATGATTTCTTTCAGGAGGCCAGACTATGATGAAACTCAGCCAT	1440
1441	Qy	 CCCAAGCTGGTTAAATTTCTATGGAGTCTGTTCAAAGGAATACCCCATATACATAGTGACT	1500
1441	Db	 CCCAAGCTGGTTAAATTTCTATGGAGTCTGTTCAAAGGAATACCCCATATACATAGTGACT	1500
1501	Qy	 GAATATATAGCAATGCTGCTCTGCTCAATTAACCTGAGGAGTACAGGAAAAGGACTTCAA	1560
1501	Db	 GAATATATAGCAATGCTGCTGCTCTGCTCAATTAACCTGAGGAGTACAGGAAAAGGACTTCAA	1560
1561	Qy	 CCTTCCAGCTCTTTAGAAATGTGCTACGATGCTGTGAAGGCATGGCCCTTCTTTGGAGAGT	1620

Db 1561 CCTTCCAGCTCTTTAGAAATGTGCTACCATGTCTGTGAAGCATGGCCCTTCTTGAGAGT 1620  
Qy 1621 CACAAATTCATACACCGGACTTGGCTGCTGTAAGTCTTGGTGGACAGATCTCTGT 1680  
Db 1621 CACAAATTCATACACCGGACTTGGCTGCTGTAAGTCTTGGTGGACAGATCTCTGT 1680  
Qy 1681 GTGAAAGTATCTGACTTTGGAAATGACAAAGTATGTTCTTGATGACCAAGTATGTCAGTTCA 1740  
Db 1681 GTGAAAGTATCTGACTTTGGAAATGACAAAGTATGTTCTTGATGACCAAGTATGTCAGTTCA 1740  
Qy 1741 GTCCGAAACAAAGTTTCCAGTCAAGTGTGACCTCCAGAGTGTTCATTTCAATATAC 1800  
Db 1741 GTCCGAAACAAAGTTTCCAGTCAAGTGTGACCTCCAGAGTGTTCATTTCAATATAC 1800  
Qy 1801 AGCAGCAAGTCAGAGCTATGGCAATTTGGGATCTGATGTGGGAGTGTTCAGCTGGG 1860  
Db 1801 AGCAGCAAGTCAGAGCTATGGCAATTTGGGATCTGATGTGGGAGTGTTCAGCTGGG 1860  
Qy 1861 AAGCAGCCCTATGACTTTGATGACAACTCCAGAGTGTTCGAAAGTCTCCAGGGCCAC 1920  
Db 1861 AAGCAGCCCTATGACTTTGATGACAACTCCAGAGTGTTCGAAAGTCTCCAGGGCCAC 1920  
Qy 1921 AGGCTTTACGGCCCTACCTGGCATCGGACACCATCTACCAAGTATGTCAGCTGTGG 1980  
Db 1921 AGGCTTTACGGCCCTACCTGGCATCGGACACCATCTACCAAGTATGTCAGCTGTGG 1980  
Qy 1981 CACGAGCTTCCAGAAACGCTCCACATTTTCCAGCACTCCCTGCTTCCATTTGAACCACTT 2040  
Db 1981 CACGAGCTTCCAGAAACGCTCCACATTTTCCAGCACTCCCTGCTTCCATTTGAACCACTT 2040  
Qy 2041 CGGAAACAAAGCAAGCAATTTGAAGAAATTTAGAGTGTCTGATGAAGTATGATATAGATG 2100  
Db 2041 CGGAAACAAAGCAAGCAATTTGAAGAAATTTAGAGTGTCTGATGAAGTATGATATAGATG 2100  
Qy 2101 CTGCGCAGCATTTTCATTCATTTTAAAGAAATTTAGAGTGTCTGATGAAGTATGATATAGATG 2160  
Db 2101 CTGCGCAGCATTTTCATTCATTTTAAAGAAATTTAGAGTGTCTGATGAAGTATGATATAGATG 2160  
Qy 2161 TTTTAAATAGTGTCTCTGATTTGCTATTTTAAAGTATGAAAGTGTCTGATGAAGTATGATATAGATG 2220  
Db 2161 TTTTAAATAGTGTCTCTGATTTGCTATTTTAAAGTATGAAAGTGTCTGATGAAGTATGATATAGATG 2220  
Qy 2221 ATTCCCTGAAATTTAGATCAAAATTTAGTAAATTTTGTATGCTCTCTGATATACAC 2280  
Db 2221 ATTCCCTGAAATTTAGATCAAAATTTAGTAAATTTTGTATGCTCTCTGATATACAC 2280  
Qy 2281 TTTCCAGCCTATAGCAGACACATTTTCAGACTGCAATATAGAGACTGTGTTTATGTTGTC 2340  
Db 2281 TTTCCAGCCTATAGCAGACACATTTTCAGACTGCAATATAGAGACTGTGTTTATGTTGTC 2340  
Qy 2341 AAAGACTGAGCAGAACTGAAATTTACTTATTTGATATTTCTTTTCTTTTATATTTGTC 2400  
Db 2341 AAAGACTGAGCAGAACTGAAATTTACTTATTTGATATTTCTTTTCTTTTATATTTGTC 2400  
Qy 2401 ATTGTCACAAATTTAATATTTACTTACCAAGTACAGAAATTTGGGAAATTTTAAACCG 2456  
Db 2401 ATTGTCACAAATTTAATATTTACTTACCAAGTACAGAAATTTGGGAAATTTTAAACCG 2456

RESULT 3  
US-10-220-801-11  
; Sequence 11, Application US/10220801  
; Publication No. US20030125235A1  
; GENERAL INFORMATION:  
; APPLICANT: FOXWELL, Brian Maurice John  
; TITLE OF INVENTION: TREATMENT OF DISEASES ASSOCIATED WITH CYTOKINE PRODUCTION WITH  
; TITLE OF INVENTION: INHIBITORS OF THE TEC FAMILY OF PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 117-412 / N95427B JP  
; CURRENT APPLICATION NUMBER: US/10/220.801  
; CURRENT FILING DATE: 2002-09-05  
; PRIOR APPLICATION NUMBER: PCT/GB01/00949  
; PRIOR FILING DATE: 2001-03-06  
; PRIOR APPLICATION NUMBER: GB 0005345.4

; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: MS Word  
; SEQ ID NO 11  
; LENGTH: 2449  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-220-801-11

Query Match 98.7%; Score 2424.4; DB 14; Length 2449;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2439; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 GCAAGCAGGAAACAGCTGAGCGGATGATATATGATACAAAATCTATTCTAGAGAA 60  
Db 1 GCAAGCAGGAAACAGCTGAGCGGATGATATATGATACAAAATCTATTCTAGAGAA 60  
Qy 61 CTTCTTCTCAAAAGATCACAGCAAAAGAAATGTCCACAAATAATTTACAAAGACGG 120  
Db 61 CTTCTTCTCAAAAGATCACAGCAAAAGAAATGTCCACAAATAATTTACAAAGACGG 120  
Qy 121 CTTTCTTCTTACCAAAACAAACCTTTCTTACTATGATATGACAAAATGAAAGGGC 180  
Db 121 CTTTCTTCTTACCAAAACAAACCTTTCTTACTATGATATGACAAAATGAAAGGGC 180  
Qy 181 AGCAGAAAGGATCCATTGAAATTAAGAAATTCAGATGTGTGGAGAAATGAAATCTCGAG 240  
Db 181 AGCAGAAAGGATCCATTGAAATTAAGAAATTCAGATGTGTGGAGAAATGAAATCTCGAG 240  
Qy 241 GAGCAGCGCTGTAGAGAGACAGTACCATTTCAGATTTCTATAAAGATGGGCTTCTC 300  
Db 241 GAGCAGCGCTGTAGAGAGACAGTACCATTTCAGATTTCTATAAAGATGGGCTTCTC 300  
Qy 301 TATGCTCTATGATCAATGAAGAGAGCGGAGTCAAGTGTGTTCAAGGATTTACAAAGAG 360  
Db 301 TATGCTCTATGATCAATGAAGAGAGCGGAGTCAAGTGTGTTCAAGGATTTACAAAGAG 360  
Qy 361 ATAAGGGTAAACCCCTGCTGCTCAAGTACCATTGTTGGTCTTCTGTCGACGGGAG 420  
Db 361 ATAAGGGTAAACCCCTGCTGCTCAAGTACCATTGTTGGTCTTCTGTCGACGGGAG 420  
Qy 421 TTTCTGTTGCTGCGAGCAGCTGTAAAGCAGCCCGAGGATGTACCTCTCGGAAAGCATAT 480  
Db 421 TTTCTGTTGCTGCGAGCAGCTGTAAAGCAGCCCGAGGATGTACCTCTCGGAAAGCATAT 480  
Qy 481 GCTAATCTGCAATCTGCAAGTCAATGAAGAGAAACACAGATTTCCACCTTCCAGACAGA 540  
Db 481 GCTAATCTGCAATCTGCAAGTCAATGAAGAGAAACACAGATTTCCACCTTCCAGACAGA 540  
Qy 541 GTGCTGAAGATACCTCGGCGAGTTCTCTTCTCAAAATGGATGACCATCTTTCAAGTACC 600  
Db 541 GTGCTGAAGATACCTCGGCGAGTTCTCTTCTCAAAATGGATGACCATCTTTCAAGTACC 600  
Qy 601 ACTCTAGCCCAATATGACAAAGAAATCTATGGTCTCCAGCCACCATCTTCA 660  
Db 601 ACTCTAGCCCAATATGACAAAGAAATCTATGGTCTCCAGCCACCATCTTCA 660  
Qy 661 AGTACCAGTCTAGCGCATATGACAGCAACTCAAGAAATCTATGGTCTCCAGCCACCAT 720  
Db 661 AGTACCAGTCTAGCGCATATGACAGCAACTCAAGAAATCTATGGTCTCCAGCCACCAT 720  
Qy 721 TTCAACATGCAATATTTCCAAAGGAGGAGTCTCCCTGACTGTGGCAAGTAAAGAACTG 780  
Db 721 TTCAACATGCAATATTTCCAAAGGAGGAGTCTCCCTGACTGTGGCAAGTAAAGAACTG 780  
Qy 781 AAAAGTAGCAGCAGTGAAGATGTTCAGAGCAAGTAAACCAAGAAAGAAATGTGAAT 840  
Db 781 AAAAGTAGCAGCAGTGAAGATGTTCAGAGCAAGTAAACCAAGAAAGAAATGTGAAT 840  
Qy 841 CACACCACCTCAAGATTTTCATGGGAATTTCCCTGAGTCAAGTTCATCTGAGCAAGAGAA 900  
Db 841 CACACCACCTCAAGATTTTCATGGGAATTTCCCTGAGTCAAGTTCATCTGAGCAAGAGAA 900

901 AACCTGGATGATATGACTGGTTGCTGGTAACATCTCCAGATCAAAATCTGAACAGTTA 960  
901 AACCTGGATGATATGACTGGTTGCTGGTAACATCTCCAGATCAAAATCTGAACAGTTA 960  
961 CTCAGACAAAGGAGGAGGAGGAGCATTTATGTTAGAAATTCGAGGCAAGTGGGAATG 1020  
961 CTCAGACAAAGGAGGAGGAGGAGCATTTATGTTAGAAATTCGAGGCAAGTGGGAATG 1020  
1021 TACACAGTGTCTTATTTAGTAAGGCTGTGAATGATATAAAGGAACTGTCAAAACATTTAC 1080  
1021 TACACAGTGTCTTATTTAGTAAGGCTGTGAATGATATAAAGGAACTGTCAAAACATTTAC 1080  
1081 CACGTGCATACAAATGCTGAGACAAATATACCTGGCAGAAACTACTGTTTGGATTC 1140  
1081 CACGTGCATACAAATGCTGAGACAAATATACCTGGCAGAAACTACTGTTTGGATTC 1140  
1141 ATTCGAAGCTTATTCATATCATCAACACAAATTCAGCAGGCATGATCACGGCTCCGC 1200  
1141 ATTCGAAGCTTATTCATATCATCAACACAAATTCAGCAGGCATGATCACGGCTCCGC 1200  
1201 CACCTGTGTCAACAAAGGCGCAACAAAGTCCCGACTCTGTGCTCCCTGGGAAATGGAATC 1260  
1201 CACCTGTGTCAACAAAGGCGCAACAAAGTCCCGACTCTGTGCTCCCTGGGAAATGGAATC 1260  
1261 TGGGAAGCTGAAAGAGAGAGATTAACCTTTGTTGAAGGAGCTGGGAAGTGGCCAGTTTGGGA 1320  
1261 TGGGAAGCTGAAAGAGAGAGATTAACCTTTGTTGAAGGAGCTGGGAAGTGGCCAGTTTGGGA 1320  
1321 GTGCTCAGCTGGGCAAGTGAAGGCGCAGTATGATGTTGCTGTTAAGATGATCAAGGAG 1380  
1321 GTGCTCAGCTGGGCAAGTGAAGGCGCAGTATGATGTTGCTGTTAAGATGATCAAGGAG 1380  
1381 GGCTCCATGTCAGAGATGATTTCTTCAGGAGCCCGCAGATGATGATGATGATGATGATGAT 1440  
1381 GGCTCCATGTCAGAGATGATTTCTTCAGGAGCCCGCAGATGATGATGATGATGATGATGAT 1440  
1441 CCCAAGCTGGTTAAATCTATGAGAGTGTTCGTAAGGAAATACCCCATATACATAGTACT 1500  
1441 CCCAAGCTGGTTAAATCTATGAGAGTGTTCGTAAGGAAATACCCCATATACATAGTACT 1500  
1501 GAATATATAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
1501 GAATATATAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
1561 CCTTCCAGCTCTTGAAGATGCTACCATCTCTGTAAGGAGTGGCCCTTCTTGGAGAGT 1620  
1561 CCTTCCAGCTCTTGAAGATGCTACCATCTCTGTAAGGAGTGGCCCTTCTTGGAGAGT 1620  
1621 CACCAATTCATACACCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
1621 CACCAATTCATACACCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
1681 GTGAAGTATGCTGCTTGGAAATGACAGAGTATGCTTGTGATGACCAAGTATGCTCAGTTCA 1740  
1681 GTGAAGTATGCTGCTTGGAAATGACAGAGTATGCTTGTGATGACCAAGTATGCTCAGTTCA 1740  
1741 GTCCGAACAAAGTTTCCAGTCAAGTGTCACTCCAGCTCCAGAGTGTTCATTTCAATAC 1800  
1741 GTCCGAACAAAGTTTCCAGTCAAGTGTCACTCCAGCTCCAGAGTGTTCATTTCAATAC 1800  
1801 AGCAGCAAGTCAAGCTGATGGGCAATTTGGGATCTCTGATGTTGGAGGTGTTCCAGCTGGGG 1860  
1801 AGCAGCAAGTCAAGCTGATGGGCAATTTGGGATCTCTGATGTTGGAGGTGTTCCAGCTGGGG 1860  
1861 AAGCAGCCCTATGACTTGTATGACAACTCCAGGTTGTTCTGAAGGTCTCCAGGGCCAC 1920  
1861 AAGCAGCCCTATGACTTGTATGACAACTCCAGGTTGTTCTGAAGGTCTCCAGGGCCAC 1920  
1921 AGGCTTTACCGGCCCCACCTGGCATTCGGACACCACTCTACCAAGATCATGATCAGCTGCTGG 1980  
1921 AGGCTTTACCGGCCCCACCTGGCATTCGGACACCACTCTACCAAGATCATGATCAGCTGCTGG 1980  
1981 CACGAGCTTCCAGAAACCGTCCCAACATTTTCAGCACTCTCTTCCATTTGAACCACTT 2040

1981 CACGAGCTTCCAGAAACCGTCCCACTTTTCAGCACTCTCTTCTCTTCCATTTGAACCACTT 2040  
2041 CGGGAAGAGCAACGATTTGAAGAAATTTAGGAGTGTCTGATAAGAAATTAATATAGATG 2100  
2041 CGGGAAGAGCAACGATTTGAAGAAATTTAGGAGTGTCTGATAAGAAATTAATATAGATG 2100  
2101 CTGGCCAGCATTTTCATTTAAGGAAATTTAGGAGTGTCTGATAAGAAATTAATATAGTGT 2160  
2101 CTGGCCAGCATTTTCATTTAAGGAAATTTAGGAGTGTCTGATAAGAAATTAATATAGTGT 2160  
2161 TTTTAATAGTGTCTCTGATTTATTTAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 2220  
2161 TTTTAATAGTGTCTCTGATTTATTTAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 2220  
2221 ATTCCCTTCAATTTAGATCAATTTAGTAAATTTGTTTATGCTGCTCTGATATATACAC 2280  
2221 ATTCCCTTCAATTTAGATCAATTTAGTAAATTTGTTTATGCTGCTCTGATATATACAC 2280  
2281 TTTCCAGCTATAGCAGAGCAGACATTTTCAGACTGCAATATAGAGACTGTGTTTATGCTGT 2340  
2281 TTTCCAGCTATAGCAGAGCAGACATTTTCAGACTGCAATATAGAGACTGTGTTTATGCTGT 2340  
2341 AAAGACTGAGCAGAACTGAAATTTACTTTATTTGGATATTCATTTCTTTTATATATGTC 2400  
2341 AAAGACTGAGCAGAACTGAAATTTACTTTATTTGGATATTCATTTCTTTTATATGTC 2400  
2401 ATTGTCAACAATTTAAATATATACCTACCAAGTACAGAAATGTGAAA 2446  
2400 ATTGTCAACAATTTAAATATATATACCTACCAAGTACAGAAATGTGAAA 2445

## RESULT 4

US-10-171-581-317  
; Sequence 317, Application US/10171581  
; Publication No. US20030104426A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: Linsley, Peter  
; APPLICANT: Mao Mao  
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia  
; FILE REFERENCE: 9301-157-999  
; CURRENT APPLICATION NUMBER: US/10/171,581  
; CURRENT FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/298,914  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 366  
; SEQ ID NO 317  
; LENGTH: 2604  
; TYPE: DNA  
; ORGANISM: Homo sapiens.  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: AF045459  
; DATABASE ENTRY DATE: 2001-06-18  
US-10-171-581-317

Query Match 97.9%; Score 2404.2; DB 14; Length 2604;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2422; Conservative 0; Mismatches 8; Indels 3; Gaps 1;

QY 21 GACGATGATAATATGGATACAAATCTATTCTAGAAGAACTTCTTCTCAAAAAGATCACA 80  
DB 175 GAAAGATGATAATATGGATACAAATCTATTCTAGAAGAACTTCTTCTCAAAAAGATCACA 234  
QY 81 GCAAAAGAGAAATGTCCCAATAATTTACAAAGAACGGCTTTTGTGTTTGACCAAAAC 140  
DB 235 GCAAAAGAGAAATGTCCCAATAATTTACAAAGAACGGCTTTTGTGTTTGACCAAAAC 294  
QY 141 AAACCTTTCTACTATGAATATGACAAAATGAAAGGGGAGCAGAAAAGGATCCATTGA 200  
DB 295 AAACCTTTCTACTATGAATATGACAAAATGAAAGGGGAGCAGAAAAGGATCCATTGA 354  
QY 201 AATTAAGAAATCAGATGCTGAGGAAACTAAATCTCGAGGAGCAGACGCCCTGTAGAGAG 260





QY 2421 TACTACCAAGTACAGAAATCTGGAAAAA 2453  
Db 2572 TACTACCAAGTACAGAAATCTGGAAAAA 2604

## RESULT 5

US-09-977-269-3  
; Sequence 3, Application US/09977269  
; Patent No. US20020082037A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHIZKY, MIKHAIL  
; APPLICANT: SURES, IRVINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1260  
; CURRENT APPLICATION NUMBER: US/09/977,269  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2500  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (82)..(2106)  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte  
; OTHER INFORMATION: kinase 2  
US-09-977-269-3

Query Match 97.6%; Score 2397.6; DB 9; Length 2500;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2417; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 22 ACGGATGATATATGGATACAAAATCTATTCTAGAGAAGCTCTCTTCAAAAAGATCACAG 81  
Db 70 AAGGATGATATATGGATACAAAATCTATTCTAGAGAAGCTCTCTTCAAAAAGATCACAG 129  
QY 82 CAAAAGAGAAATGTCACCAAAATTAATACAAAGAACGGCTTTTGTGACCAAAACA 141  
Db 130 CAAAAGAGAAATGTCACCAAAATTAATACAAAGAACGGCTTTTGTGACCAAAACA 189  
QY 142 AACCTTTCTCTACTATGATATGACAAAATGAAAGGGGCGAGCAAAAGGATCCATTGAA 201  
Db 190 AACCTTTCTCTACTATGATATGACAAAATGAAAGGGGCGAGCAAAAGGATCCATTGAA 249  
QY 202 ATTAAGAAATCAGATGTGTGGAGAAAGTAAATCTCAGGAGCAGAGCGCTGTAGAGAGA 261  
Db 250 ATTAAGAAATCAGATGTGTGGAGAAAGTAAATCTCAGGAGCAGAGCGCTGTAGAGAGA 309  
QY 262 CAGTACCATTTCAGATGTCTATAGATGGCTCTCTATGCTATGATCATCAATGAA 321  
Db 310 CAGTACCATTTCAGATGTCTATAGATGGCTCTCTATGCTATGATCATCAATGAA 369  
QY 322 GAGAGCGGAAGTCAAGTGGTGTCTTCTGTCGACGGGAAGTAAAGGGTAAACCCCACTG 381  
Db 370 GAGAGCGGAAGTCAAGTGGTGTCTTCTGTCGACGGGAAGTAAAGGGTAAACCCCACTG 429  
QY 382 CTGGTCAAGPACCATAGTGGGTCTTCTGTCGACGGGAAGTAAAGGGTAAACCCCACTG 441  
Db 430 CTGGTCAAGPACCATAGTGGGTCTTCTGTCGACGGGAAGTAAAGGGTAAACCCCACTG 489  
QY 442 TGTAAAGCAGCCAGGATGATACCTCTGGAGAGCATATGCTAATTCGATACGATGCAATG 501  
Db 490 TGTAAAGCAGCCAGGATGATACCTCTGGAGAGCATATGCTAATTCGATACGATGCAATG 549  
QY 502 AATGAAGAGAAACACAGAGTTCACCTCTCCAGAGAGTGTCTGAAGATACCTCGGGCA 561  
Db 550 AATGAAGAGAAACACAGAGTTCACCTCTCCAGAGAGTGTCTGAAGATACCTCGGGCA 609

QY 562 GTTCCTGTCTTCAAAAATGGATGACCAATCTTCAAGTACCACCTCTAGCCCAATATGACAAC 621  
Db 610 GTTCCTGTCTTCAAAAATGGATGACCAATCTTCAAGTACCACCTCTAGCCCAATATGACAAC 669  
QY 622 GAATCAAAAGAAAACTATGGCTCCAGCCACCATCTTCAAGTACCAGTCTAGCGCAATAT 681  
Db 670 GAATCAAAAGAAAACTATGGCTCCAGCCACCATCTTCAAGTACCAGTCTAGCGCAATAT 729  
QY 682 GACAGCAACTCAAAAGAAATCTATGGCTCCAGCCCAATCTTCAAGTACCAGTCTATATTC 741  
Db 730 GACAGCAACTCAAAAGAAATCTATGGCTCCAGCCCAATCTTCAAGTACCAGTCTATATTC 789  
QY 742 AGGGAAGACTTCCCTGACTGGTGGCAAGTAAGAAACTGAAAAGTACAGCAGCAGTGA 801  
Db 790 AGGGAAGACTTCCCTGACTGGTGGCAAGTAAGAAACTGAAAAGTACAGCAGCAGTGA 849  
QY 802 GATGTTGCAAGCAGTAACCAAAAAGAAAGTGAATACACACCACTCAAAAGATTTC 861  
Db 850 GATGTTGCAAGCAGTAACCAAAAAGAAAGTGAATACACACCACTCAAAAGATTTC 909  
QY 862 TGGGAATTCCTGAGTCAAGTTCACTCTGAAGAGAGGAAACCTGGATGATTATGACTGG 921  
Db 910 TGGGAATTCCTGAGTCAAGTTCACTCTGAAGAGAGGAAACCTGGATGATTATGACTGG 969  
QY 922 TTTGCTGTAACATCTCCAGATCACAATCTGACAGTTACTCAGACAAAGGAAAGAA 981  
Db 970 TTTGCTGTAACATCTCCAGATCACAATCTGACAGTTACTCAGACAAAGGAAAGAA 1029  
QY 982 GGACATTTATGTTAGAAATTCGAGCAAGTGGCAATGTACACAGTGTCTTATTTAGT 1041  
Db 1030 GGACATTTATGTTAGAAATTCGAGCAAGTGGCAATGTACACAGTGTCTTATTTAGT 1089  
QY 1042 AAGCTGTGAATGATAAAAAAGAACTGTCAAACTATACCAGTGTCCATACAAATGCTGAG 1101  
Db 1090 AAGCTGTGAATGATAAAAAAGAACTGTCAAACTATACCAGTGTCCATACAAATGCTGAG 1149  
QY 1102 AACAAATTTATACCTGGCAGAAACTTACTGTTTGTGATTCATTCCTCAAAAGCTTAT 1161  
Db 1150 AACAAATTTATACCTGGCAGAAACTTACTGTTTGTGATTCATTCCTCAAAAGCTTAT 1209  
QY 1162 CATCAACAAATTCAGCAGGATGATCACAGGCTCCGACCCCTGTGTCAAAAGAGGC 1221  
Db 1210 CATCAACAAATTCAGCAGGATGATCACAGGCTCCGACCCCTGTGTCAAAAGAGGC 1269  
QY 1222 AACAGGTCCCGACTCTGTCTCCCTGGGAAATGGAATCTGGGAACTGGAAGAGAGAG 1281  
Db 1270 AACAGGTCCCGACTCTGTCTCCCTGGGAAATGGAATCTGGGAACTGGAAGAGAGAG 1329  
QY 1282 ATTACCTTTGAAGGAGCTGGGAAAGTGGCAAGTTGGAGTGTCCAGCTGGGCAAGTGG 1341  
Db 1330 ATTACCTTTGAAGGAGCTGGGAAAGTGGCAAGTTGGAGTGTCCAGCTGGGCAAGTGG 1389  
QY 1342 AAGGGCAGATGATGTTGCTGTTAAGATGATCAAGAGAGGCTCCATGTCAAGAGATGA 1401  
Db 1390 AAGGGCAGATGATGTTGCTGTTAAGATGATCAAGAGAGGCTCCATGTCAAGAGATGA 1449  
QY 1402 TTCTTTTCAGGAGGCCAGACTATGATGAAACTCAGCCATCCCAAGCTGGTAAATTCAT 1461  
Db 1450 TTCTTTTCAGGAGGCCAGACTATGATGAAACTCAGCCATCCCAAGCTGGTAAATTCAT 1509  
QY 1462 GGAGTGTGTTCAAAGGAATACCCCATATACATAGTACTGATATATAGCAATGGCTGC 1521  
Db 1510 GGAGTGTGTTCAAAGGAATACCCCATATACATAGTACTGATATATAGCAATGGCTGC 1569  
QY 1522 TTGCTGAATTAACCTGAGAGTACCGGAAAGAGACTTGAACCTTCCAGCTCTTACAATG 1581  
Db 1570 TTGCTGAATTAACCTGAGAGTACCGGAAAGAGACTTGAACCTTCCAGCTCTTACAATG 1629  
QY 1582 TGCTACGATGTCTGTAAGGATGAGCCCTTCTTGGAGAGTCAACAAATTCATACACCGGAC 1641  
Db 1630 TGCTACGATGTCTGTAAGGATGAGCCCTTCTTGGAGAGTCAACAAATTCATACACCGGAC 1689



|||||  
Db 910 TGGGATTCCTGAGTCAAGTTTCATCTGAAGAGAGGAAACCTGGATGATTATGACTGG 969  
Qy 922 TTTGCTGGTAACATCTCCAGATCACAATCTGAACAGTTACTCAGACAAAGGAAAGAA 981  
Db 970 TTTGCTGGTAACATCTCCAGATCACAATCTGAACAGTTACTCAGACAAAGGAAAGAA 1029  
Qy 982 GGAGCATTTATGGTTAGAAATTCAGCCCAAGTGGGAATGTACACAGTGTCTTATTTAGT 1041  
Db 1030 GGAGCATTTATGGTTAGAAATTCAGCCCAAGTGGGAATGTACACAGTGTCTTATTTAGT 1089  
Qy 1042 AAGGCTGTGAATGATAAAGGAAAGTGTCAAAACATTTACCAGTGTCAACAAATGCTGAG 1101  
Db 1090 AAGGCTGTGAATGATAAAGGAAAGTGTCAAAACATTTACCAGTGTCAACAAATGCTGAG 1149  
Qy 1102 AACAAATTTATACCTGGCAGAAATCTGTTTTGATTCATTCATTCCTTATTCATTTAT 1161  
Db 1150 AACAAATTTATACCTGGCAGAAATCTGTTTTGATTCATTCATTCCTTATTCATTTAT 1209  
Qy 1162 CATCAACAAATTCAGCAGGATGATCAGCGCTCCGCCACCCCTGTGTCAACAAGGCC 1221  
Db 1210 CATCAACAAATTCAGCAGGATGATCAGCGCTCCGCCACCCCTGTGTCAACAAGGCC 1269  
Qy 1222 AACAGGTCCTCGACTCTGTCTCCCTGGGAAATGGAATCTGGGAATGAAAGAGAGAG 1281  
Db 1270 AACAGGTCCTCGACTCTGTCTCCCTGGGAAATGGAATCTGGGAATGAAAGAGAGAG 1329  
Qy 1282 ATTACCTTTGTAAGAGAGCTGGGAAGTGGCCAGTTTGGAGTGTCCAGCTGGCGAAGTGG 1341  
Db 1330 ATTACCTTTGTAAGAGAGCTGGGAAGTGGCCAGTTTGGAGTGTCCAGCTGGCGAAGTGG 1389  
Qy 1342 AAGGGCAGTATGATGTGCTGTTAAGATGATCAAGGAGGCTCCATGTCAGAGATGAA 1401  
Db 1390 AAGGGCAGTATGATGTGCTGTTAAGATGATCAAGGAGGCTCCATGTCAGAGATGAA 1449  
Qy 1402 TTTCTTTAGGAGGCCAGACTATGATGAACCTCAGCCATCCCAAGCTGGTTAAATCTAT 1461  
Db 1450 TTTCTTTAGGAGGCCAGACTATGATGAACCTCAGCCATCCCAAGCTGGTTAAATCTAT 1509  
Qy 1462 GGAGTGTGTTCAAAGGAATACCCATATACATAGTACTGACTGAATATATAAGCAATGGCTGC 1521  
Db 1510 GGAGTGTGTTCAAAGGAATACCCATATACATAGTACTGACTGAATATATAAGCAATGGCTGC 1569  
Qy 1522 TTGCTGAATACCTGAGAGTCAAGGAAAGGACTTGAACCTTCCAGCTCTTGAAGATG 1581  
Db 1570 TTGCTGAATACCTGAGAGTCAAGGAAAGGACTTGAACCTTCCAGCTCTTGAAGATG 1629  
Qy 1582 TGCTACGATGCTGTGAAGGATGGCCCTTCTTGGAGAGTCAACCAATTCATACACGGGAC 1641  
Db 1630 TGCTACGATGCTGTGAAGGATGGCCCTTCTTGGAGAGTCAACCAATTCATACACGGGAC 1689  
Qy 1642 TTGCTGCTCCTAACTGCTGTGGTGGACAGATCTCTGTGAAAGTATCTGACTTTGGA 1701  
Db 1690 TTGCTGCTCCTAACTGCTGTGGTGGACAGATCTCTGTGAAAGTATCTGACTTTGGA 1749  
Qy 1702 ATGACAGGATGTTCTTTGATGACCAAGTATGTCAGTTCAGTGGGAAACAAAGTTTCCAGTC 1761  
Db 1750 ATGACAGGATGTTCTTTGATGACCAAGTATGTCAGTTCAGTGGGAAACAAAGTTTCCAGTC 1809  
Qy 1762 AAGTGGTCAGTCCAGAGGCTTTTCAATTAATACAGCAGCAAGTCAAGCATATGG 1821  
Db 1810 AAGTGGTCAGTCCAGAGGCTTTTCAATTAATACAGCAGCAAGTCAAGCATATGG 1869  
Qy 1822 GCATTTGGGATCTGATGTGGAGGTGTTTCAGCCTGGGAAAGCAGCCCTATGACTTGTAT 1881  
Db 1870 GCATTTGGGATCTGATGTGGAGGTGTTTCAGCCTGGGAAAGCAGCCCTATGACTTGTAT 1929  
Qy 1882 GACAACTCCCAAGTGGTTCTGAAGGTCTCCAGGCGCCACAGGCTTTACCGGCCCCACCTG 1941  
Db 1930 GACAACTCCCAAGTGGTTCTGAAGGTCTCCAGGCGCCACAGGCTTTACCGGCCCCACCTG 1989  
Qy 1942 GCATCGACACCATCTACCATGATCATGTACAGTGTGCTGGCAGAGCTTTCCAGAAAGCGT 2001  
|||||

Db 1990 GCATCGGACACCATCTACAGATCATGTACAGCTGCTGGCACGAGCTTCCAGAAAGCCGT 2049  
Qy 2002 CCCACATTTTCAGCAACTCCTGCTTCCATTGAACCACTTCCGGGAAAGAAAGCAATTTGA 2061  
Db 2050 CCCACATTTTCAGCAACTCCTGCTTCCATTGAACCACTTCCGGGAAAGAAAGCAATTTGA 2109  
Qy 2062 AGAAGAAATTTAGAGTGTGCTGATGAAGTGAATAGATGCTGCCAGCATTTTCATTTAT 2121  
Db 2110 AGAAGAAATTTAGAGTGTGCTGATGAAGTGAATAGATGCTGCCAGCATTTTCATTTAT 2169  
Qy 2122 TTTAAGGAAATTTAGAGTGTGCTGATGAAGTGAATTTAGTGTGTTTAAATAGTGTCTCTGAT 2181  
Db 2170 TTTAAGGAAATTTAGAGTGTGCTGATGAAGTGAATTTAGTGTGTTTAAATAGTGTCTCTGAT 2229  
Qy 2182 TGTCTATTATTAGAAATGAACAAGGAGGAAACAAAGATTCCTTTGAAATTTAGATCA 2241  
Db 2230 TGTCTATTATTAGAAATGAACAAGGAGGAAACAAAGATTCCTTTGAAATTTAGATCA 2289  
Qy 2242 AATTAGTAAATTTTGTATGCTGCTGCTGATATACACTTTCCAGCCTATAGCAGAGC 2301  
Db 2290 AATTAGTAAATTTTGTATGCTGCTGCTGATATACACTTTCCAGCCTATAGCAGAGC 2348  
Qy 2302 ACATTTTCAGACTGCAATATAGAGACTGTGTTTCATGTGTAAGACTGAGCAGAACTGAAA 2361  
Db 2349 ACATTTTCAGACTGCAATATAGAGACTGTGTTTCATGTGTAAGACTGAGCAGAACTGAAA 2408  
Qy 2362 AATTACTATTGGAATATTCATTTCTTTTATATTATTTGTCATGTCACAAATTAATAT 2421  
Db 2409 AATTACTATTGGAATATTCATTTCTTTTATATTATTTGTCATGTCACAAATTAATAT 2468  
Qy 2422 ACTACCAAGTACAGAAATGTGGAAAAA 2453  
Db 2469 ACTACCAAGTACAGAAATGTGGAAAAA 2500

RESULT 7  
US-09-977-261-3  
; Sequence 3, Application US/09977261  
; Publication No. US20030054527A1  
; GENERAL INFORMATION:  
; APPLICANT: GISHIZKY, MIKHAIL  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: SURES, IRMLINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1259  
; CURRENT APPLICATION NUMBER: US/09/977,261  
; PRIOR FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2500  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (82)..(2106)  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte  
; OTHER INFORMATION: Kinase 2  
US-09-977-261-3

Query Match 97.6%; Score 2397.6; DB 11; Length 2500;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2417; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

Qy 22 ACGATGATATATATGATACAAATCTATTCTAGAAAGTCTTCTCTCAAGATCACAG 81  
Db 70 AAGGATGATATATGATACAAATCTATTCTAGAAAGTCTTCTCTCAAGATCACAG 129  
Qy 82 CAAAAGAGAAATGTACCAAAATATTACAAAGAGCGGCTTTGTTGTACCAAAACA 141  
|||||

Qy 22 ACGATGATATATATGATACAAATCTATTCTAGAAAGTCTTCTCTCAAGATCACAG 81  
Db 70 AAGGATGATATATGATACAAATCTATTCTAGAAAGTCTTCTCTCAAGATCACAG 129  
Qy 82 CAAAAGAGAAATGTACCAAAATATTACAAAGAGCGGCTTTGTTGTACCAAAACA 141  
|||||

Db 130 CAAAGAAGAAATGTCAACCAATATATACAAAGAACGGCTTTTGTGTTTGACCAAAACA 189  
-QY 142 AACCTTCCCTACTATGATATGACAAATGAAAGGGGCGACGAAAGGATCCATTGAA 201  
Db 190 AACCTTCCCTACTATGATATGACAAATGAAAGGGGCGACGAAAGGATCCATTGAA 249  
QY 202 ATTAAGAAATACAGATGTGTGGAAAGTAAATCTCGAGGAGCAGAGCCCTGTAGAGAGA 261  
Db 250 ATTAAGAAATACAGATGTGTGGAAAGTAAATCTCGAGGAGCAGAGCCCTGTAGAGAGA 309  
QY 262 CAGTACCAATTCAGATGTCTATAAGATGGGCTCTCTATGCTATGCAATCAATGAA 321  
Db 310 CAGTACCAATTCAGATGTCTATAAGATGGGCTCTCTATGCTATGCAATCAATGAA 369  
QY 322 GAGAGCGAAGTCAGTGTGAAAGCATACAAAGAGATAAGGGGTAAACCCCACTG 381  
Db 370 GAGAGCGAAGTCAGTGTGAAAGCATACAAAGAGATAAGGGGTAAACCCCACTG 429  
QY 382 CTGCTCAAGTACCATGTGGGTTCTCGTGACGGGAAGTTCTCTGTGTGGCAGCAGAGC 441  
Db 430 CTGCTCAAGTACCATGTGGGTTCTCGTGACGGGAAGTTCTCTGTGTGGCAGCAGAGC 489  
QY 442 TGTAAAGCAGCCAGCATGTACCTCTGGGAAGCATATGCTAATCTGCATACGACATC 501  
Db 490 TGTAAAGCAGCCAGCATGTACCTCTGGGAAGCATATGCTAATCTGCATACGACATC 549  
QY 502 AATGAAGAGAAACACAGAGTCCACCTTCCAGACAGAGTGTGGAAGATACCTCGGGCA 561  
Db 550 AATGAAGAGAAACACAGAGTCCACCTTCCAGACAGAGTGTGGAAGATACCTCGGGCA 609  
QY 562 GTTCCCTGTTCTCAAAATGATGACCATCTTCAAGTACCAGCTTAGGCCCAATATGACAAC 621  
Db 610 GTTCCCTGTTCTCAAAATGATGACCATCTTCAAGTACCAGCTTAGGCCCAATAT 669  
QY 622 GAATCAAGAAATACATAGGTCGCCAGCACCATCTTCAAGTACCAGCTTAGGCCCAATAT 681  
Db 670 GAATCAAGAAATACATAGGTCGCCAGCACCATCTTCAAGTACCAGCTTAGGCCCAATAT 729  
QY 682 GACAGCAACCAAGAAATCTATGGCTCCAGCAGCAACCTTCAACATGCACTATATCCA 741  
Db 730 GACAGCAACCAAGAAATCTATGGCTCCAGCAGCAACCTTCAACATGCACTATATCCA 789  
QY 742 AGGGAAGACTTCCCTGACTGTGGCAAGTAAAGAACTGAAAGTAGCAGCAGAGTGA 801  
Db 790 AGGGAAGACTTCCCTGACTGTGGCAAGTAAAGAACTGAAAGTAGCAGCAGAGTGA 849  
QY 802 GATGTTCCAGCAGTAAACCAAGAAAGAAATGTGAAATCACACCCTCAAGATTTC 861  
Db 850 GATGTTCCAGCAGTAAACCAAGAAAGAAATGTGAAATCACACCCTCAAGATTTC 909  
QY 862 TGGGAATTCCTGAGTCAAGTTCATCTGAAGAGAGAAACCTGGGATGATTATGACTGG 921  
Db 910 TGGGAATTCCTGAGTCAAGTTCATCTGAAGAGAGAAACCTGGGATGATTATGACTGG 969  
QY 922 TTTGCTGGTAACATCTCCAGATCAATCTGAACAGTTACTGACAGAAAAGGGAAGAA 981  
Db 970 TTTGCTGGTAACATCTCCAGATCAATCTGAACAGTTACTGACAGAAAAGGGAAGAA 1029  
QY 982 GGAGCATTTATGTTAGAAATTCGAGCAAGTGGGAATGTACAGAGTCTCTATTATAGT 1041  
Db 1030 GGAGCATTTATGTTAGAAATTCGAGCAAGTGGGAATGTACAGAGTCTCTATTATAGT 1089  
QY 1042 AAGCTGTGAATGATAAAAAGGAACGTCAACACATTACCAGTGGCATACAAATGCTGAG 1101  
Db 1090 AAGCTGTGAATGATAAAAAGGAACGTCAACACATTACCAGTGGCATACAAATGCTGAG 1149  
QY 1102 AACAAATATACCTGGCAGAAAACCTACTGTTTGAATTCCTATTCCTCAAGAGCTTATTCATTAT 1161  
Db 1150 AACAAATATACCTGGCAGAAAACCTACTGTTTGAATTCCTATTCCTCAAGAGCTTATTCATTAT 1209  
QY 1162 CATCAACAAATTCAGCAGGCATGATCACAGGTCGCCGACCCCTGTGTCAACAAGGCC 1221  
Db 1210 CATCAACAAATTCAGCAGGCATGATCACAGGTCGCCGACCCCTGTGTCAACAAGGCC 1269

QY 1222 AACAGGTCCCGACTGTGTCTCCCTGGGAAATGGAATCTGGAACTGAAAAGAGAAGAG 1281  
Db 1270 AACAGGTCCCGACTGTGTCTCCCTGGGAAATGGAATCTGGAACTGAAAAGAGAAGAG 1329  
QY 1282 ATTACCTTTGTTAAAGGAGCTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGG 1341  
Db 1330 ATTACCTTTGTTAAAGGAGCTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGG 1389  
QY 1342 AAGGGCAGTATGATGTGCTGTTAAGATGATCAAGGAGGCTCCATGTCAGAAATGAA 1401  
Db 1390 AAGGGCAGTATGATGTGCTGTTAAGATGATCAAGGAGGCTCCATGTCAGAAATGAA 1449  
QY 1402 TTCTTTTCAGGAGGCCAGACTATGATGAAACTCAGCCATCCCAAGCTGGTTAAATTCAT 1461  
Db 1450 TTCTTTTCAGGAGGCCAGACTATGATGAAACTCAGCCATCCCAAGCTGGTTAAATTCAT 1509  
QY 1462 GGAGTGTGTTCAAAAGGAATACCCATATACATAGTACTGAATATATAGCAATGGCTGC 1521  
Db 1510 GGAGTGTGTTCAAAAGGAATACCCATATACATAGTACTGAATATATAGCAATGGCTGC 1569  
QY 1522 TTGCTGAATACCTGAGGAGTCACGGAAGAGACTTGAACCTTCCAGCTCTTAGAAATG 1581  
Db 1570 TTGCTGAATACCTGAGGAGTCACGGAAGAGACTTGAACCTTCCAGCTCTTAGAAATG 1629  
QY 1582 TGCTACGATGTCTGTGAAGGATGGCTTCTTGGAGAGTCAACCAATTCATACACCGGAC 1641  
Db 1630 TGCTACGATGTCTGTGAAGGATGGCTTCTTGGAGAGTCAACCAATTCATACACCGGAC 1689  
QY 1642 TTGCTGCTCTAACTGCTTGGTGACAGATCTCTGTGGAAGTATCTGACTTTGGA 1701  
Db 1690 TTGCTGCTCTAACTGCTTGGTGACAGATCTCTGTGGAAGTATCTGACTTTGGA 1749  
QY 1702 ATGACAGGTATGTTCTTGATGACCACTATGTCAGTTCAGTGGGACAAAGTTTCCAGTC 1761  
Db 1750 ATGACAGGTATGTTCTTGATGACCACTATGTCAGTTCAGTGGGACAAAGTTTCCAGTC 1809  
QY 1762 AAGTGGTCAGTCCAGAGGTGTTTCATTACTTCAATACAGCAAGTCAAGCAAGTATGG 1821  
Db 1810 AAGTGGTCAGTCCAGAGGTGTTTCATTACTTCAATACAGCAAGTCAAGCAAGTATGG 1869  
QY 1822 GCATTTGGGATTCCTGATGTGGGAGGTGTCAGCCTGGGGAAGCAGCCCTATGACTTGTAT 1881  
Db 1870 GCATTTGGGATTCCTGATGTGGGAGGTGTCAGCCTGGGGAAGCAGCCCTATGACTTGTAT 1929  
QY 1882 GACAACTCCCAAGTGGTCTGAAAGTCTCCAGGCCCCACAGCTTTACCGGCCCCCAGCTG 1941  
Db 1930 GACAACTCCCAAGTGGTGGTCTGAAAGTCTCCAGGCCCCACAGCTTTACCGGCCCCCAGCTG 1989  
QY 1942 GCATCGGACACCATCTACAGATCATGTACAGTCTGCGCAGAGCTTCCAGAAAGCT 2001  
Db 1990 GCATCGGACACCATCTACAGATCATGTACAGTCTGCGCAGAGCTTCCAGAAAGCT 2049  
QY 2002 CCCACATTTACAGCAACTCTCTCTTCATTTGAACTTTCAGGGAAGGAAAGCAAGCAATGGA 2061  
Db 2050 CCCACATTTACAGCAACTCTCTCTTCATTTGAACTTTCAGGGAAGGAAAGCAAGCAATGGA 2109  
QY 2062 AGAAGAAATAGGAGTCTGATAGAAATGATATAGATGCTGGCCAGCAATTTTCATTCAT 2121  
Db 2110 AGAAGAAATAGGAGTCTGATAGAAATGATATAGATGCTGGCCAGCAATTTTCATTCAT 2169  
QY 2122 TTTAAGGAAATAGGAGTCTGATAGAAATGATATAGTACTGTTTAAAGTGTCTCTGTAT 2181  
Db 2170 TTTAAGGAAATAGGAGTCTGATAGAAATGATATAGTACTGTTTAAAGTGTCTCTGTAT 2229  
QY 2182 TGTCTATTATTAGAAATGAAAGGAGGAGGAAACAAAAGATTCCTCTGAAATTTAGATCA 2241  
Db 2230 TGTCTATTATTAGAAATGAAAGGAGGAGGAAACAAAAGATTCCTCTGAAATTTAGGTC 2289  
QY 2242 AATTAGTAAATTTGTTTATGCTGCTGATATAGCACTTTCAGGCTTATAGCAGAGC 2301  
Db 2290 AATTAGTAAATTTGCTGCTGATATAGCACTTTCAGGCTTATAGCAGAGC 2348





Qy	1187	TCACACGGCTCCGCCACCCCTGTGTCAACAAGGCCCAAGTCCCGCACTCTGTGTGTCC	1246
Db	1241	TATCCAGCGTGAARATATCCTGTGTCTAAACAAACAAACAGCGCCCTTCTACTGCAAGCC	1300
Qy	1247	TGSGAAATGGAATCTCGGAACTGAAAGAGAGAGAGTACCTTGTCTGAAGAGCTGGAA	1306
Db	1301	TGGGCTATGAGTCAATGGGAATATGATCCAAAGGACCTCACCTCTTGAAGAGGCTGGGA	1360
Qy	1307	GTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGAAGGGGCAGTATGATGTTGCTGTTA	1366
Db	1361	CTGSACAATTCGGTGTCTGTAATATGSGAAGTGGAGGGCCAAATATGATGTGGCCATCA	1420
Qy	1367	AGATGATCAGGAGGGCTCATGTCAAGAGATGAATTCCTTAGGAGGCCAGACTATGA	1426
Db	1421	AGATGATCAGAGAAGTTTCCATGTCGAGGATGAATTCATTGAAGAAGCCAAAGTCTATGA	1480
Qy	1427	TGAAATCAGCACTCCCAAGCTGGTTAAATTCATGAGGTGTGTTCAAAGGAATACCCCA	1486
Db	1481	TGAATCTTCCCATGAGAAGCTGGTGCAGTTGTATGGGTCGTCACCAACACAGCCCA	1540
Qy	1487	TATACATAGTGAAGTAAATATATAGCAATGGCTGCTGTGTAATTAACCTGAGGAGTCAG	1546
Db	1541	TCTTCATCATCACCGAGTACATGGCTAATGGCTCCCTTGAACCTACCTGAGGAGATGC	1600
Qy	1547	GAAAGAGCTTGAACCTTCCAGCTCTTAGAATGTGCTAGGATGTCTGTGAAGGCATGG	1606
Db	1601	GGCACCGCTTCCAGACACACAGCACTGCTTGAGATGTGCAAGATGCTGTGAAGCAATGG	1660
Qy	1607	CCTTCTTTGGAGAGTCAACCAATTCATACACGGGACTTGGTGTCTGTAACTGCTGTGTGG	1666
Db	1661	AATACTTTGGAGTCCGAGCGATTCTCTACAGAGACCTGGCAGCTCGGAAACTGTTTGGTAA	1720
Qy	1667	ACAGAGATCTCTGTGAAAGTATCTGACTTTTGGAAATGACAAGGTATGTTCTTGATGACC	1726
Db	1721	ACGATCAAGGAGTTGTGAAAGTATCTGACTTTTGGCCCTGTCTAGTATGTCTCTGATCATG	1780
Qy	1727	AGTATGTCAGTTCAGTCCGGAACAAGTTTCCAGTCAAGTGGTCAAGTCCAGAGGTGTTTC	1786
Db	1781	AGTACACCAAGCTCTGTAGGCTCCAAAGTTTCCAGTCCGGTGGTCTCCACCAAGAAGTGCTTA	1840
Qy	1787	ATTACTTCAAAATACAGACGAAGTCAAGCACTATGGGCATTTGGGATCCTGATGTGGCAGG	1846
Db	1841	TGTATACAAAGTTCAGCAGCAAAATCTGCATCTGGGCTTTTGGGGTTTAAATGTGGGAGA	1900
Qy	1847	TGTTCAAGCTGGGGAAGACGCCCTATGACTTTGATATGACAACCTCCCAAGTGGTTCCTGAAGG	1906
Db	1901	TCTACTCCCTGGGGAAGATGCGGTATGAGAGATTTACTTAACAGTGAAGACAGCAGAACA	1960
Qy	1907	TCTCCCAAGGCCACAGCGTTTACCGGCCCACTGGCATCGGACACCATCTACACAGATCA	1966
Db	1961	TTGCTCAAGGCTTACGTCCTTACAGGCGCTCATCTGGCATCAGAGGGGTATATACCATCA	2020
Qy	1967	TGTACAGTGTGTCAGAGGCTTCCAGAAAAGCGTCCACATTTTCAGCAACTCCTGTCTTT	2026
Db	2021	TGTACAGCTGCTGCACAGAGAAACGATGACCTCTAGTTTCAAAAATCTCTTGTAGTA	2080
Qy	2027	CCATTGAACCACTTCGGGAAAAGA	2051
Db	2081	ACATTCATAGTGTGATGTGAAGA	2105

## RESULT 10

RESULT TO  
US-10-045-202-1

US-10-043-202-1  
: Sequence 1, Application US/10045202; sequence 1, Application US/1004  
; Publication NO. US20030040461A1

; FUDIFICATION NO. 0320  
; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb

APPLICANT: Bristol Myers Squibb

; AFFRICANI: BISCOT-Myers squids  
; TITLE OF INVENTION: MODULATORS OF

; TITLE OF INVENTION: INTERMEDIATES AND METHODS FOR THEIR IDENTIFICATION AND USE IN THE  
; DATE OF INVENTION: MODIFICATIONS OF PROTON 3 LYSINE KINASE AND PROTON 3 LYSINE KINASE  
; TITLE OF INVENTION: INTERMEDIATES AND METHODS FOR THEIR IDENTIFICATION AND USE IN THE

; TITLE OF INVENTION: PREVENTION OF OSTEOPOROSIS AND RELATED DISEASE STATES

FILE OF INVENTION: 1  
; FILE REFERENCE: D0032 NP

```

; CURRENT APPLICATION NUMBER: US/10/045,202
;
; CURRENT FILING DATE: 2002-06-27
;
; PRIOR APPLICATION NUMBER: 60/242,471
;
; PRIOR FILING DATE: 2000-10-23
;
; NUMBER OF SEQ ID NOS: 12
;
; SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 1
;
; LENGTH: 2560
;
; TYPE: DNA
;
; ORGANISM: Human
;
US-10-045-202-1

```

Query Match

Best Local Similarity 63.58; Pred. No. 1.9e-119;

Best Local Similarity	03.5%	Frequency	1.56	Indels	0	Gaps	0
Matches	740	Conservative	0	Mismatches	425		

QY	887	CTGAAGAGAGGAAACCTCGATGATTTATGACTCGTGGTAAACATCTCCAGATCAC	946
DB	941	CTTGAAGAGAGAGACTCCATAGAAATGTTAGTGGTATTTCCAAACACATGACTCGGAGTC	1000
QY	947	AATCTGAACAGTTACTCAGACAAAGGAAAAGAGAGCATTTATGGTTAGAAATTCGA	1006
DB	1001	AGGCTGAGCAACTGCTAAAGCAGAGGGAAAGAGGGTTTCATTTGCAGACACTCCA	1060
QY	1007	GCCAAAGTGGGAATGTACACAGTGCCTTTATTAAGAGGCTGTGAAATGATAAAAAGGAA	1066
DB	1061	GCAAAGCTGGCAAAATATACAGTGTCTGTGTTTGTCTAATCCACAGGGGACCCCTCAAGGG	1120
QY	1067	CTGTCAACATTTACCAGTGCATACAATGCTGAGAACAAATTAATCTGGSCACAAAACCT	1126
DB	1121	TGATACGTCAATATGTTGTGTGTTCCACACTCAGAGCCAGTATTAATCTGGCTGAGAGC	1180
QY	1127	ACTGTTTGTATTCATTCCAAAGCTTATTCATTAATCATCAACACAATTCAGCAGGCGATGA	1186
DB	1181	ACCTTTTCAGCACCATCCCTCAGCTCATTAATTAATACCATCAGCACAACTCTCAGGACTCA	1240
QY	1187	TCACACGGCTCCGCCACCTGTGTCAACAAGGCCACACAGGTCCTCCGACTCTGTGTCCC	1246
DB	1241	TATCCAGGCTCCAAATATCCAGTGTCTCAACAAAACAAGATGCACTTCCACTGCAGGCC	1300
QY	1247	TGGAAATGGAACTCTGGGAATGAAAGAGAGAAGATTAACCTTGTGAAGGAGCTGGGAA	1306
DB	1301	TGGGATACGGATCATGGGAATTCATCCAAAGGACCTCACCTTCTGAAGGAGCTGGGA	1360
QY	1307	GTGGCAGTTTGGAGTGTCCAGCTGGCAAGTGGAGGGCAGTATGATGTTGCTGTTA	1366
DB	1361	CTGGACAAATTTGGGGTAGTGAAGTATGGGAAATGGAGGGCCAGTAGCACTGGCCATCA	1420
QY	1367	AGATGATCAAGGAGGGCTCCATGTCAAGAAGATGAATTTTCAGAGGCCCCAGACTATGA	1426
DB	1421	AGATGATCAAGAGAGGCTCCATGTCTGNAGATGAATTCATGAGAGGCCAAGATCATGA	1480
QY	1427	TGAACACTAGCCATCCCAAGCTGGTTAAATTTCTATGGAGTGTGTTCAAGGAATACCCCA	1486
DB	1481	TGAATCTTTCCCATGAGAAGCTGGTCAGATTGTATGGCGTCTGCACCAAGCAGGCCCCA	1540
QY	1487	TATACATAGTACGATATATAGCAATGCTGCTGTCTCAATTAACCTGAGGAGTCAAG	1546
DB	1541	TCTTCATCATCAGTAGTACATGCCAAATGGCTGCCTCTGAACTACCTGAGGGAGATGC	1600
QY	1547	GAAAGGAGCTTGAACCTTTCCAGCTCTTGAATAATGTGTACGATCTCTGTGAAGGCATGG	1606
DB	1601	GCCACCGTTCCAGNACTCAGCAGCTGCTAGAGATGCAAGATGCTCTGTGAAGCCATGG	1660
QY	1607	CCTTCTTGGAGAGTCAACAAATTCATACACCGGGACTTGGCGTGCCTGTAACCTTGGTGG	1666
DB	1661	AATACCTGGAGTCAAAAGCAGTTCCCTTCACCGAGAGCTGGCAGCTCGAAACTGTTTGGTAA	1720
QY	1667	ACAGAGATCTCTGTGTGAAGTATCTGACTTTTGGAAATCACAAGGTATGTTCTTCATGACC	1726
DB	1721	ACGATCAAGGAGTTGTTTAAAGATCTGTGATTTTCGGCTCTGCAGGTTATGTCTCTGATGATG	1780
QY	1727	AGTATGTCAGTTCAGTCCGGAACAAAGTTTCCAGTCAAGTGTGTGCTCAGCTCCAGAGGTGTTTC	1786

1781 AATACACAGCTCAGTAGCTCAATTTCCAGTCGGTGGTCCCAACGGAGTCTCTGA 1840  
1787 ATTACTTCAATACAGCAGCAGTATGAGTATGGGATCTTGGATCTCTGATGGAGG 1846  
1841 TGTATGACAAAGTTCACAGCAATCTGACATTTGGGCTTTGGGGTTTGTGATGTTGGGAAA 1900  
1847 TGTTCAGCTGGGGAAGCAGCCCTATGACTTGTATGACAACTCCCAAGGTGTTCTGAAGG 1906  
1901 TTTACTCCCTGGGGAAGATGCCATATGAGAGATTACTAACAGTGAGACTGCTGAACACA 1960  
1907 TCTCCAGGCCCCACAGGCTTTACCGGCCCCACTGGCATCGGACACCATCTACAGATCA 1966  
1961 TTGCCCAAGGCTACGCTCTACAGGCTCATCTGGCTTCAGAGAGGTATATACCATCA 2020  
1967 TGTACAGCTGTCGACAGCTTCAGAAAAGCGTCCCAATTTTCAGCAACTCTCTGCTTT 2026  
2021 TGTACAGTTGTTGGCATGAGAAAGCAGATGAGCGTCCCACTTTCAAAAATTTCTCTGAGCA 2080  
2027 CCATTGAACCACTTCGGGAAAAGA 2051  
2081 ATATTCTAGATGTCATGGATGAAGA 2105

## RESULT 11

US-10-220-801-4

; Sequence 4, Application US/10220801

; Publication No. US20030125235A1

; GENERAL INFORMATION:

; APPLICANT: FOXWELL, Brian Maurice John

; TITLE OF INVENTION: TREATMENT OF DISEASES ASSOCIATED WITH CYTOKINE PRODUCTION WITH

; TITLE OF INVENTION: INHIBITORS OF THE TEC FAMILY OF PROTEIN TYROSINE KINASES

; FILE REFERENCE: 117-412 / N83427B JP

; CURRENT APPLICATION NUMBER: US/10/220,801

; PRIOR FILING DATE: 2002-09-05

; PRIOR APPLICATION NUMBER: PCT/GB01/00949

; PRIOR FILING DATE: 2001-03-06

; PRIOR APPLICATION NUMBER: GB 0005345.4

; PRIOR FILING DATE: 2000-03-06

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: MS Word

; SEQ ID NO 4

; LENGTH: 2582

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-220-801-4

Query Match 19.6%; Score 480.2; DB 14; Length 2582;

Best Local Similarity 63.3%; Pred. No. 3.8e-118;

Matches 737; Conservative 0; Mismatches 428; Indels 0; Gaps 0;

887 CTGAAGAAGAGAAACCTGGATGATATGACTGGTGGTAAACATCTCCAGATCAC 946

972 CTGAAGCAGAGACTCCATAGAAATGTATGATGCTATCCAAACACATGACTCGGAGTC 1031

947 AATCTGAACAGTTACTCAGACAAAGGAAAGAGAGCAATTTATGTTAGAAATTCGA 1006

1032 AGGCTGAGCAACTGCTAAGCAAGAGGAAAGAGAGGTTTCATTGTCAGAGACTCCA 1091

1007 GCCAAGTGGGAATGTACACAGTCTCTTATTAGTAAAGCTGTGAATGNTAAAAAGGNA 1066

1092 GCAAGCTGGCAATATACAGTGTCTGTTTGTCTAAATCCACAGGGGACCTCCAGGGG 1151

1067 CTGTCAACATATACACGTGTCATACAAATGTGTGAGAACAAATATACCTGGCAGAAAACT 1126

1152 TGATACGTCATATGTTGTGTCTCCACACCTCAGAGCCAGTATTACCTGGCTGAGAGC 1211

1127 ACTGTTTGTATCCCAATCCCAAGCTTATTCATATATCAACAAATTCAGCAGCATCA 1186

1212 ACCTTTTTCAGCAGCACCCTCCTGAGCTCATTAATACCATCAGCAGCAACTCTGAGGACTCA 1271

1187 TCACAGGCTCGGCCACCTGTGTCAACAAGGCAAGGTCGCCGACTCTGTGTCTCC 1246

1272 TATCCAGGCTCAATATCCAGTCTCTCAACAAAACAAAGATGACCTTCCACTTCGACGCC 1331  
1247 TGGGAATATGGAATCTGGGAATGAAAGAGAGAGATTAACCTTGTGTTGAAGAGCTGGAA 1306  
1332 TGGATACGGATCATGGGAAATGATCCAAAGGACCTGACCTTCTTGAAGGAGCTGGGA 1391  
1307 GTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGGGAAGGGCCAGTATGATGTTGCTCTTA 1366  
1392 CTGGCAATTTTGGGAGTAGTCAAGTATGGGAAATGGAGAGGCCAGTACGACGTGGCCATCA 1451  
1367 AGATGATCAAGAGGCTCCATGTCAGAGATCAATTTCTTTCAGGAGGCCAGACTATGA 1426  
1452 AGATGATCAAGAGGCTCCATGTCGAGAGATCAATTCATTTGAGAGCCAAAGTCAATGA 1511  
1427 TGAATCTCAGCCATCCCAAGCTGGTTAAATTTCTTTCAGGAGTGTGTTTCAAGAAATACCCCA 1486  
1512 TGAATCTTTCATCATGAGAGCTGGTGCAGTTGTATGGCGTCTGCACCAAGCAGCGCCCA 1571  
1487 TATACATAGTACTGATATATATAGCAATGGGCTGCTTCTGTAATTAACCTCAGGAGTACG 1546  
1572 TCTTCATCATCACTGAGTACATGGCAATGGCTCCCTCTGAACTACCTCAGGAGATGC 1631  
1547 GAAAGGACTTGAACCTTCCAGCTCTTAGAAATGTGCTACGATGCTGTGAAGGCAATGG 1606  
1632 GCCACGGCTTCAGACTCAGCAGCTGCTAGAGATGTGCAAGGATGCTGTGAAGCCATGG 1691  
1607 CCTTCTTGGAGAGTACCAATTCATACACCGGAGCTTGGCTGCTCTCACTGCTTGTGTTGG 1666  
1692 AATACCTGAGTCAAGAGCTTCTTCCACCGAGACCTGGCAGCTCGAAACTGTTTGGTAA 1751  
1667 ACAGAGATCTCTGTGTAAGATATGACTTTTGAATGACAGAGTATGTTCTTGTATGAC 1726  
1752 AGCATCAAGAGAGTGTGTAAGATATCTGATTTCCGCGCTGTCCAGGATATGCTCTGATGATG 1811  
1727 AGTATCTCAGTTTCAGTCCGGAACAAAGTTTCCAGTCAAGTGGTCCAGCTCCAGAGGTGTTTC 1786  
1812 AATACAAAGCTCAGTAGGCTCCAAATTTCCAGTCCGGTGGTCCCAAGGAGTCTCTGA 1871  
1787 ATTACTTCAATACAGCAGCAGTATGAGCTATGGGCAATTTGGGATCTCTGATGTTGGAGG 1846  
1872 TGTATAGCAAGTTTACAGCAAAATCTGACATTTGGGCTTTGGGGTTTTCATGTTGGGAAA 1931  
1847 TCTTACGCTGGGGAAGCAGCCCTATGACTTGTATGACAACTCCCAAGGTGTTCTGAAGG 1906  
1932 TTTACTCCCTGGGGAAGATGCCATATGAGAGATTTACTAACAGTGCAGCTGCTGAACACA 1991  
1907 TCTCCAGGCCCACAGGCTTTACCGGCCCCACTGCGCATCGGACACCATCTACAGATCA 1966  
1992 TTGCCCAAGGCTACGCTCTACAGGCTCATCTGGCTTCAGAGAGGTTATATACCATCA 2051  
1967 TGTACAGCTGCTGGCAGCAGCTTCCAGAAAAGCGTCCCAATTTTCAGCAACTCTCTGCTTT 2026  
2052 TGTACAGTTGTTGGTTTATAGAAAGCAGATGAGCGTCCCACTTTTCAAAATTTCTTCTGAGCA 2111  
2027 CCATTGAACCACTTCGGGAAAAGA 2051  
2112 ATATTCTAGATGTCATGGATGAAGA 2136

## RESULT 12

US-09-733-103-2

; Sequence 2, Application US/09735103

; Patent No. US20010005589A1

; GENERAL INFORMATION:

; APPLICANT: Mano, Hiroyuki

; APPLICANT: Sakata, Tsunekaki

; APPLICANT: Hasegawa, Mamoru

; TITLE OF INVENTION: Promoter

; FILE REFERENCE: 50026/011002

; CURRENT APPLICATION NUMBER: US/09/735,103

; CURRENT FILING DATE: 2000-12-12

; PRIOR APPLICATION NUMBER: 09/142,529

; PRIOR FILING DATE: 1998-09-09







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2003, 11:40:22 ; Search time 628 Seconds

(without alignments)  
10557.038 Million cell updates/sec

Title: US-10-021-660-41

Perfect score: 2456

Sequence: 1 gcaagcagcggaacagctga.....aatgtggagaaaaaaaccg 2456

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum First 40%

Listing first 15 summaries

Database :

N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2456	100.0	2456	25	ABX93297
2	2456	100.0	2456	25	ACAS6878
3	2456	100.0	2456	25	ABX08803
4	2424.4	98.7	2449	22	AAS14050
5	2388.6	97.3	2494	16	AAT00617
6	494.6	20.1	2468	24	ABN3858
7	493	20.1	2505	17	AAT45648
8	485	19.7	1365	24	ABK85707

9	485	19.7	2560	24	ABN83857	Human Bruton's tyr
10	480.2	19.6	2582	22	AAS14047	Human cDNA encoding
11	437.8	17.8	2574	18	ABT73125	Mouse Tec tyrosine
12	394.4	16.1	3663	24	ABT62789	Prostate adenocarc
13	392.8	16.0	3650	23	AAS70719	DNA encoding novel
14	388	15.8	3593	22	AAS14048	Human cDNA encoding
15	385.4	15.7	1368	24	ABK85703	DNA encoding trunc
16	385.4	15.7	4221	22	AAS14049	Human cDNA encoding
17	385.4	15.7	6383	24	ABL63135	Breast cancer rela
18	385.4	15.7	6401	23	AAS80103	DNA encoding novel
19	383.8	15.6	6381	24	ABV94054	Breast carcinoma r
20	292.4	11.9	3236	23	ABL13689	Drosophila melanog
21	263.4	10.7	431	22	ABH34243	Human colon cancer
22	237.6	9.7	20155	23	ABL13688	Drosophila melanog
23	226.8	9.2	3653	24	ABT99918	Rat mucocuticular ce
24	226.8	9.2	4705	24	ABA96146	Modified human/mou
25	226.8	9.2	4786	24	ABA96147	Modified human/mou
26	223.4	9.1	3323	25	ABZ18519	Group III cDNA can
27	223.4	9.1	3380	25	ABZ18520	Group III cDNA can
28	223.4	9.1	5520	18	ABT61865	c-abl gene. Homo
29	223.4	9.1	5527	24	ABN97215	Gene #3713 used to
30	220.2	9.0	3623	19	AAV20457	Human c-abl oncoge
31	220.2	9.0	3780	18	AAT61864	c-abl gene. Homo
32	220.2	9.0	3840	25	AAD52781	Human Abl DNA. Ho
33	220.2	9.0	5434	24	ABV77964	Hypoxia-repressed
34	211.6	8.6	1614	25	ABX74390	Human cDNA sequenc
35	211.6	8.6	4414	24	ABZ35401	Human gene express
36	207.6	8.5	1090	23	AAS87964	DNA encoding novel
37	207.6	8.5	1418	17	AAT45649	Human haematopoiet
38	207.6	8.5	1611	14	AAQ46688	Human pp60 c-src g
39	207.6	8.5	1611	24	ABT73325	DNA encoding human
40	207.6	8.5	1611	25	ABZ59382	DNA encoding novel
41	207.6	8.5	1699	23	AAS87963	Human cDNA sequenc
42	207.6	8.5	3299	22	AAH18556	Novel human coding
43	207.6	8.5	4466	24	ABN59752	cDNA sequence enco
44	206	8.4	3849	25	ABX10962	Wild-type chicken
45	205.6	8.4	1759	21	AAZ29700	

#### ALIGNMENTS

##### RESULT 1

ABX93297  
ID ABX93297 standard; cDNA; 2456 BP.

XX AC ABX93297;

XX DT 04-JUN-2003 (first entry)

XX DE cDNA encoding human bone marrow tyrosine kinase gene in chromosome X.

XX KW Human; vascular endothelium; Bmx TK; arteriosclerosis; endocardial cell;  
KW bone marrow tyrosine kinase gene in chromosome X; arteriosclerosis;  
KW arterial endothelial cell; growth signal; smooth muscle.  
KW inflammatory response; non-thrombogenic property; antiinflammatory;  
KW Bmx tyrosine kinase signalling pathway; antiarteriosclerotic;  
KW chromosome Xp22.2; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 34..2061

XX FT /\*tag= a

XX FT /product= "Bmx TK"

XX US2002173481-A1.

XX PD 21-NOV-2002.

XX PF 01-JUL-2002; 2002US-0186399.

XX XX 25-JUN-1998; 98US-0104863.





```
QY 1561 CCTTCCAGCTCTTAGAAATGCTGTACATGCTGTGAAGCATGGCCCTTCTTGAGAGT 1620
Db 1561 CCTTCCAGCTCTTAGAAATGCTGTACATGCTGTGAAGCATGGCCCTTCTTGAGAGT 1620
QY 1621 CACCAATTCATACACCGGACTTGGCTGCTGTAACCTGCTGTGGGACAGAGATCTCTGT 1680
Db 1621 CACCAATTCATACACCGGACTTGGCTGCTGTAACCTGCTGTGGGACAGAGATCTCTGT 1680
QY 1681 GTGAAAGTATCTGACTTTGGAATGACAAGGTATGTTCTTGATGACCAAGTATGTCAGTTCA 1740
Db 1681 GTGAAAGTATCTGACTTTGGAATGACAAGGTATGTTCTTGATGACCAAGTATGTCAGTTCA 1740
QY 1741 GTGGAAACAAGTTTCCAGTCAAGTGTGTCAGCTCCAGAGGTGTTTCAATTAATCAATAC 1800
Db 1741 GTGGAAACAAGTTTCCAGTCAAGTGTGTCAGCTCCAGAGGTGTTTCAATTAATCAATAC 1800
QY 1801 AGCAGCAAGTCAGACGTATGGCATTTGGATGCTGATGCTGGAGGTGTTTCAGCCTGGGG 1860
Db 1801 AGCAGCAAGTCAGACGTATGGCATTTGGATGCTGATGCTGGAGGTGTTTCAGCCTGGGG 1860
QY 1861 AAGCAGCCCTATGACTTGTATGACAACCTCCAGGTGTTCTGAAAGGTCTCCAGGGCCAC 1920
Db 1861 AAGCAGCCCTATGACTTGTATGACAACCTCCAGGTGTTCTGAAAGGTCTCCAGGGCCAC 1920
QY 1921 AGCCTTTACGGGCCACCTGGCATCGGACACCATCTACAGATCATGTATGATGCTGTGG 1980
Db 1921 AGCCTTTACGGGCCACCTGGCATCGGACACCATCTACAGATCATGTATGATGCTGTGG 1980
QY 1981 CAGCAGCTTCCAGAAAGCGTCCACATTTCCAGCACTCTGCTCTTCCATTTGAACCACTT 2040
Db 1981 CAGCAGCTTCCAGAAAGCGTCCACATTTCCAGCACTCTGCTCTTCCATTTGAACCACTT 2040
QY 2041 CGGGAAACACAGCAATTCAGGAAGAAATAGSAGTCTGATAGAAATGAATATAGATG 2100
Db 2041 CGGGAAACACAGCAATTCAGGAAGAAATAGSAGTCTGATAGAAATGAATATAGATG 2100
QY 2101 CTGGCAGCATTTTCATCAATTTAGGAAGTAGGAAGCATTAAGTAATTTAGCTAGT 2160
Db 2101 CTGGCAGCATTTTCATCAATTTAGGAAGTAGGAAGCATTAAGTAATTTAGCTAGT 2160
QY 2161 TTTTATAGTGTCTCTGATTTGCTATTATTTAGAAATGAACAGCAGCAACAAAG 2220
Db 2161 TTTTATAGTGTCTCTGATTTGCTATTATTTAGAAATGAACAGCAGCAACAAAG 2220
QY 2221 ATTCCCTTGAATTTAGATCAAAATTTAGTAATTTGTTTATGCTGCTCTCATATAACAC 2280
Db 2221 ATTCCCTTGAATTTAGATCAAAATTTAGTAATTTGTTTATGCTGCTCTCATATAACAC 2280
QY 2281 TTTCCAGCCTATAGCAGAACACATTTTCAGACTGCAATATAGAGACTGTGTTCAATGCT 2340
Db 2281 TTTCCAGCCTATAGCAGAACACATTTTCAGACTGCAATATAGAGACTGTGTTCAATGCT 2340
QY 2341 AAAGACTGAGCAGCACTGAAATTTACTTATTTGATATTTCTTCTTTTATATTTGTC 2400
Db 2341 AAAGACTGAGCAGCACTGAAATTTACTTATTTGATATTTCTTCTTTTATATTTGTC 2400
QY 2401 ATTGTCAACAATTAATATATCTACCAAGTACAGAAATGTTGGAATTTTATTTTATTTGTC 2456
Db 2401 ATTGTCAACAATTAATATATCTACCAAGTACAGAAATGTTGGAATTTTATTTTATTTGTC 2456
```

## RESULT 2

ACA56878

ID ACA56878 standard; cDNA; 2456 BP.

XX

AC ACA56878;

XX

DT 06-JUN-2003 (first entry)

XX

DE Human signalling pathway polynucleotide probe SEQ ID NO 1476.

XX

KW Human; probe; ss; array element; Parkinson's disease;

KW signalling pathway population; cancer; adenocarcinoma; leukaemia;  
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.  
XX Homo sapiens.  
XX US500938-B1.  
XX 31-DEC-2002.  
XX 30-JAN-1998; 98US-0016434.  
XX 30-JAN-1998; 98US-0016434.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Au-Young J, Seilhamer JJ;  
XX WPI; 2003-352189/33.  
XX  
XX Combination of polynucleotide probes, useful as array elements in a  
XX microarray for monitoring the expression of a number of target  
XX polynucleotides -  
XX  
XX Claim 1; SEQ ID NO 1476; 65pp; English.

The invention relates to a combination which, comprises a number of polynucleotide probes comprising a sequence selected from one of the 1490 sequences mentioned in the specification. The combination is useful as an array element in a microarray for monitoring the expression of a number of target polynucleotides. The microarray is particularly useful in the diagnosis and treatment of cancer and immunopathology and neuropathology. The microarray is useful in diagnostics and treatment regimens, drug discovery and development, toxicological and carcinogenicity studies, forensics and pharmacogenomics. The microarray is also useful for monitoring progression of diseases and for developing sophisticated profiles for the effects of currently available therapeutic drugs. The combination is also useful for purifying a subpopulation of mRNAs, cDNAs and genomic fragments and in research and diagnostic applications. The array can detect changes in expression in a large number of genes coding for different signaling pathway populations which can be used to diagnose various diseases including cancer e.g. adenocarcinoma and leukaemia, immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease and Parkinson's disease. The present sequence represents a polynucleotide probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docID=06500938B1.

QY Sequence 2456 BP; 805 A; 495 C; 549 G; 607 T; 0 other;

Query Match 100.0%; Score 2456; DB 25; Length 2456;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GCAAGCAGGAAACGAGTGTGAGCGGATGATATATGATACAAATCTATTCTAGAAGAA 60
Db 1 GCAAGCAGGAAACGAGTGTGAGCGGATGATATATGATACAAATCTATTCTAGAAGAA 60
QY 61 CTTCCTTCTCAAAAGATCACAGCAAAAGAAATGTCACCAATATTAATACAAAGCGG 120
Db 61 CTTCCTTCTCAAAAGATCACAGCAAAAGAAATGTCACCAATATTAATACAAAGCGG 120
QY 121 CTTTCTTCTTGTGACCAAAACAACTTCTCTTACTATGATATGACAAATGAAAAGGGC 180
Db 121 CTTTCTTCTTGTGACCAAAACAACTTCTCTTACTATGATATGACAAATGAAAAGGGC 180
QY 181 AGCAGAAAAGATTCATTGAAATTAAGAAAATCAGATGTGGAGAAAAGTAAATCTCGAG 240
Db 181 AGCAGAAAAGATTCATTGAAATTAAGAAAATCAGATGTGGAGAAAAGTAAATCTCGAG 240
QY 241 GAGCAGACGCTGTAGAGACACAGTACCCATTTTCTAGATTTCTATAGAGATGGCTTCTC 300
Db 241 GAGCAGACGCTGTAGAGACACAGTACCCATTTTCTAGATTTCTATAGAGATGGCTTCTC 300
```





1201 CACCCCTGTGTCAAAAGGCCCAACAAAGGTCGCCGACTCTGTGTCTCCCTGGGAATGGAATC 1260  
1201 CACCCCTGTGTCAAAAGGCCCAACAAAGGTCGCCGACTCTGTGTCTCCCTGGGAATGGAATC 1260  
1261 TGGGAAGCTGAAAGAGAGAGATACCTCTGTGTGAAGGAGCTGGGAAGTGCCAGTTTGGGA 1320  
1261 TGGGAAGCTGAAAGAGAGAGATACCTCTGTGTGAAGGAGCTGGGAAGTGCCAGTTTGGGA 1320  
1321 GTGGTCCAGCTGGCAAGTGGGAAGGGGCGAGTATGATGTCTGTGAAGTATCAAGGAG 1380  
1321 GTGGTCCAGCTGGCAAGTGGGAAGGGGCGAGTATGATGTCTGTGAAGTATCAAGGAG 1380  
1381 GGTCTCATCTCAGAGATGAATCTCTTTCAGGAGGCCGAGACTATGATGAACATCAGCCAT 1440  
1381 GGTCTCATCTCAGAGATGAATCTCTTTCAGGAGGCCGAGACTATGATGAACATCAGCCAT 1440  
1441 CCCAAGCTGGTTAAATCTTATGGAGTGTCTTCAAGGAAATACCCCATATACATAGTACT 1500  
1441 CCCAAGCTGGTTAAATCTTATGGAGTGTCTTCAAGGAAATACCCCATATACATAGTACT 1500  
1501 GAATATATAGCAATGGCTGCTGTCTGAATACCTCAGAGTACCGGAAAGGACTTGAA 1560  
1501 GAATATATAGCAATGGCTGCTGTCTGAATACCTCAGAGTACCGGAAAGGACTTGAA 1560  
1561 CTTTCCAGCTCTTACAAATGTCTACGATGTCTGTGAAGGATGGCTTCTTGGAGAGT 1620  
1561 CTTTCCAGCTCTTACAAATGTCTACGATGTCTGTGAAGGATGGCTTCTTGGAGAGT 1620  
1621 CACCAATTCATACACGGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
1621 CACCAATTCATACACGGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
1681 GTGAAGTATCAGCTTTGGAATGCAAGGATATGTTCTTGTATGACCAAGTATGCACTTCA 1740  
1681 GTGAAGTATCAGCTTTGGAATGCAAGGATATGTTCTTGTATGACCAAGTATGCACTTCA 1740  
1741 GTCCGAACAAAGTTTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
1741 GTCCGAACAAAGTTTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
1801 AGCAGCAAGTCAAGCTATGGGATTTGGGATCCTGATGTGGGAGTGTTCAGCCCTGGGG 1860  
1801 AGCAGCAAGTCAAGCTATGGGATTTGGGATCCTGATGTGGGAGTGTTCAGCCCTGGGG 1860  
1861 AGCAGCCCTATGACTTTGATGCAACTCCAGTGTCTGAGTGTCTCCAGGCGCCAC 1920  
1861 AGCAGCCCTATGACTTTGATGCAACTCCAGTGTCTGAGTGTCTCCAGGCGCCAC 1920  
1921 AGGCTTTTACCGGCCCCACCTGGCATCGGACACCATCTACAGATCATGTACAGCTGCTGG 1980  
1921 AGGCTTTTACCGGCCCCACCTGGCATCGGACACCATCTACAGATCATGTACAGCTGCTGG 1980  
1981 CAGGAGCTTCCAGAAAGCGTCCACATTTACAGCAACTCCTGTCTCTCCATTTGAACCACTT 2040  
1981 CAGGAGCTTCCAGAAAGCGTCCACATTTACAGCAACTCCTGTCTCTCCATTTGAACCACTT 2040  
2041 CGGGAAGAGACAGCAATTCAGCAAGAAATAGGAGTGTCTGATGAAGAAATATAGATG 2100  
2041 CGGGAAGAGACAGCAATTCAGCAAGAAATAGGAGTGTCTGATGAAGAAATATAGATG 2100  
2101 CTGGCCAGCATTTTCATTTTAAAGAAAGTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2160  
2101 CTGGCCAGCATTTTCATTTTAAAGAAAGTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2160  
2161 TTTTATAGTGTCTCTGTATGCTTATTTATTTAGAAATGAACAAGGAGGAGGAAACAAAG 2220  
2161 TTTTATAGTGTCTCTGTATGCTTATTTATTTAGAAATGAACAAGGAGGAGGAAACAAAG 2220  
2221 ATTCCTTGAATTTAGATCAAAATAGTATTTTGTGTTTATGCTGCTCTGTATATACAC 2280  
2221 ATTCCTTGAATTTAGATCAAAATAGTATTTTGTGTTTATGCTGCTCTGTATATACAC 2280  
2281 TTTCAGCCTATAGCAGAGCACATTTTCAGACTGCAATATAGAGACTGTGTCTATGTGT 2340

2281 TTTCCAGCTTATAGCAGACACATTTTCAGACTGCAATATAGAGACTGTGTCTATGTGT 2340  
2341 AAGACTGACGACAACTGAAATATTACTTATTGGATATTACTTCTTTTATATATGTC 2400  
2341 AAGACTGACGACAACTGAAATATTACTTATTGGATATTACTTCTTTTATATATGTC 2400  
2401 ATTGTCAACAATTAATATATCTACCAAGTACAGAAATGGAAGGAAAAAAGCCG 2456  
2401 ATTGTCAACAATTAATATATCTACCAAGTACAGAAATGGAAGGAAAAAAGCCG 2456

RESULT 4  
AAS14050  
ID AAS14050 standard; cDNA; 2449 BP.  
XX  
AC AAS14050;  
XX AC  
XX AC  
DT 18-DEC-2001 (first entry)  
XX  
DE Human cDNA encoding partial protein tyrosine kinase, Bmx.  
XX  
KW Immunosuppressive; antiinflammatory; osteopathic; antirheumatic;  
KW antiarthritic; dermatological; antipsoriatic; antispasmodic; cytostatic;  
KW cardiant; antiarteriosclerotic; vasotropic; cerebroprotective;  
KW neuroprotective; nontropic; ophthalmological; immunosuppressive;  
KW hepatotropic; virucide; anti-HIV; immunomodulator; antithyroid; sepsis;  
KW septic shock; inflammation; Crohn's disease; rheumatoid arthritis;  
KW leukopenia; osteoarthritis; spondyloarthropathy; psoriatic arthritis;  
KW severe steroid resistant asthma; pulmonary fibrosis; myocarditis;  
KW atherosclerosis; angina; vasculitis; reperfusion injury; infarction;  
KW stroke; multiple sclerosis; Alzheimer's disease; Graves ophthalmopathy;  
KW psoriasis; transplant rejection; acute alcoholic hepatitis; fibrosis;  
KW HIV; human immunodeficiency virus; cachexia; erythema nodosum lepromatous;  
KW borreliosis; meningococcal septicaemia; cancer; human; Bmx; ss;  
KW antibacterial; protein tyrosine kinase.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
CDS 1..2061  
FT /\*tag= a  
FT /product= "Bmx"  
FT /partial  
FT /note= "No start codon"  
XX  
PN WO200166107-A2.  
XX  
XX 13-SEP-2001.  
XX  
XX 06-MAR-2001; 2001WO-GB00949.  
XX  
XX 06-MAR-2000; 2000GB-0005345.  
XX  
XX (KENN-) KENNEDY INST RHEUMATOLOGY MATHILDA & TER.  
XX  
XX Foxwell BMJ;  
XX  
XX WPI; 2001-602596/68.  
XX P-PSDB; AAU09013.  
XX  
XX Treating condition associated with cytokine production in a mammal,  
XX e.g. sepsis, septic shock, inflammation and Crohn's disease, comprises  
XX administering an inhibitor of a member(s) of the Tec family of protein  
XX tyrosine kinases  
XX  
XX Claim 23; Fig 18; 78pp; English.  
XX  
XX The invention relates to the use of inhibitors of the Tec family of  
XX protein tyrosine kinases to treat a condition associated with cytokine  
XX production in a mammal. The Tec family members are e.g. Bruton's  
XX tyrosine kinase (Btk), Tec, Itk, Txk or Bmx. The condition is especially  
XX sepsis, septic shock, inflammation, Crohn's disease or a condition

induced by TRR (Toll related receptor) ligand, LPS (lipopolysaccharide),  
 CC zymosan or Gram-negative bacteria. The method may also be used for  
 CC treating rheumatoid arthritis. Biological effects of LPS include fever,  
 CC circulatory disturbances and vascular hypersensitivity to adrenergic  
 CC drugs, leukopenia typically followed by leucocytosis, non-specific  
 CC stimulation of B-lymphocytes to undergo blast transformation and  
 CC proliferation, lethal toxicity and non-specific tolerance to endotoxin  
 CC through repeated exposure to LPS. LPS also causes an increase in  
 CC production of cytokines e.g. TNF. Conditions associated with IL-1beta  
 CC and TNFalpha osteoarthritis, spondyloarthropathy, psoriatic arthritis,  
 CC severe steroid resistant asthma, pulmonary fibrosis, myocarditis,  
 CC atherosclerosis, angina with high CRP/IL-6, vasculitis, reperfusion  
 CC injury, infarction, stroke, multiple sclerosis, Alzheimer's disease,  
 CC Graves ophthalmopathy, psoriasis, transplant rejection, acute alcoholic  
 CC hepatitis, fibrosis. Other conditions for which inhibitors of TLR  
 CC family PTKs may be used include HIV (human immunodeficiency virus  
 CC infection), cachexia, erythema nodosum leprostatum, borreliosis, or  
 CC meningococcal septicemia, cancer, such as breast cancer, ovarian cancer  
 CC or colon cancer. Many other examples of diseases are given in the  
 CC specification. The present sequence encodes human Bmx.  
 XX  
 QQ Sequence 2449 BP; 797 A; 495 C; 548 G; 609 T; 0 other;

Query Match 98.7%; Score 2424.4; DB 22; Length 2449;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 2439; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY	1	GCAAGCAGGAAACAGCTGAGCGGATGATATATGGATACAAATCTATCTCAAGAA	50
DB	1	GCAAGCAGGAAACAGCTGAGCGGATGATATATGGATACAAATCTATCTCAAGAA	50
QY	61	CTTCTTCTCAAAAGATCACAGCAAAAGAAAGAAATGTCCACCAATAATATACAAAGCGG	120
DB	61	CTTCTTCTCAAAAGATCACAGCAAAAGAAAGAAATGTCCACCAATAATATACAAAGCGG	120
QY	121	CTTTTGTGTTTACCAAAACAAACCTTTCTACTATGAATATGACAAATGAAAGGGGCG	180
DB	121	CTTTTGTGTTTACCAAAACAAACCTTTCTACTATGAATATGACAAATGAAAGGGGCG	180
QY	181	AGCAGAAAGATCCATTGAATTAAGAAATGAGATGATGAGGAGAAATGATCTCGAG	240
DB	181	AGCAGAAAGATCCATTGAATTAAGAAATGAGATGATGAGGAGAAATGATCTCGAG	240
QY	241	GAGCAGAGCGCTGTAGAGAGACAGTACCATTTCAGATTGCTATAAAGATGGGCTCTC	300
DB	241	GAGCAGAGCGCTGTAGAGAGACAGTACCATTTCAGATTGCTATAAAGATGGGCTCTC	300
QY	301	TATGCTATGATCAATGAAGAGAGCGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	360
DB	301	TATGCTATGATCAATGAAGAGAGCGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	360
QY	361	ATAAGGGTATACCCGACCTGCTGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	420
DB	361	ATAAGGGTATACCCGACCTGCTGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	420
QY	421	TTCCTGTCTTCCAGCAGAGCTGTAAAGCAGCCCGGAGTGTACCTCTGGGAACATAT	480
DB	421	TTCCTGTCTTCCAGCAGAGCTGTAAAGCAGCCCGGAGTGTACCTCTGGGAACATAT	480
QY	481	GCTAATCTGCATCTGCAAGTCAATGAAGAGAGAAACACAGTTCACACCTTCCAGACAGA	540
DB	481	GCTAATCTGCATCTGCAAGTCAATGAAGAGAGAAACACAGTTCACACCTTCCAGACAGA	540
QY	541	GTGCTGAAGATACCTCGGGAGTTCCTGTTCTCAAAATGGATGACACCTCTTCAAGTACC	600
DB	541	GTGCTGAAGATACCTCGGGAGTTCCTGTTCTCAAAATGGATGACACCTCTTCAAGTACC	600
QY	601	ACTTAGCCCAATATGACAAAGATCAAGAAAGAAACTATGGTCCCGACCACTCTTCA	660
DB	601	ACTTAGCCCAATATGACAAAGATCAAGAAAGAAACTATGGTCCCGACCACTCTTCA	660
QY	661	AGTACAGTCTAGCGCAATATGACAGCACTCAAGAAATCTATGGTCCCGACCAAC	720
DB	661	AGTACAGTCTAGCGCAATATGACAGCACTCAAGAAATCTATGGTCCCGACCAAC	720

DB	661	AGTACAGTCTAGCGCAATATGACAGCACTCAAGAAATCTATGGTCCCGACCAAC	720
QY	721	TTCAACATGAGTATATTCACAGGAGAGTTCCTCTGACGTGGGCAAGTAAAGAACTG	780
DB	721	TTCAACATGAGTATATTCACAGGAGAGTTCCTCTGACGTGGGCAAGTAAAGAACTG	780
QY	781	AAAAGTAGCAGCAGCAGTGAAGATGTTGCAAGCAGTAACCAAAAGAAAGAAATGTCAAT	840
DB	781	AAAAGTAGCAGCAGCAGTGAAGATGTTGCAAGCAGTAACCAAAAGAAAGAAATGTCAAT	840
QY	841	CACACACCTCAAGATTTTCATGGGAATTCCTCTGAGTCAAGTTCATCTGAAGAGAGAA	900
DB	841	CACACACCTCAAGATTTTCATGGGAATTCCTCTGAGTCAAGTTCATCTGAAGAGAGAA	900
QY	901	AACTGGATGATATGACGTGGTGGTAAATCTCCAGATCAACATCTGAACAGTTA	960
DB	901	AACTGGATGATATGACGTGGTGGTAAATCTCCAGATCAACATCTGAACAGTTA	960
QY	961	CTCAGACAAAAGGAAAGAGAGCATTTATGTTAGAAATTCGAGCCCAAGTGGGAATG	1020
DB	961	CTCAGACAAAAGGAAAGAGAGCATTTATGTTAGAAATTCGAGCCCAAGTGGGAATG	1020
QY	1021	TACACAGTCTCTTATTTAGTAAAGCTGGAATGATAAAAGAACTGTCAACATAC	1080
DB	1021	TACACAGTCTCTTATTTAGTAAAGCTGGAATGATAAAAGAACTGTCAACATAC	1080
QY	1081	CAGCTGCATACAAATGCTGAGAACAAATATACCTGGCAGAAACTACTCTTTGATTC	1140
DB	1081	CAGCTGCATACAAATGCTGAGAACAAATATACCTGGCAGAAACTACTCTTTGATTC	1140
QY	1141	ATTCCAAAGCTTATTCATTAATCAACAAATTCAGCAGCATGATCAGCAGCTCCGC	1200
DB	1141	ATTCCAAAGCTTATTCATTAATCAACAAATTCAGCAGCATGATCAGCAGCTCCGC	1200
QY	1201	CACCTGTGTCAACAAAGCCCAACAGTCCCGACTCTGTCTCTCTGGGAAATGGAATC	1260
DB	1201	CACCTGTGTCAACAAAGCCCAACAGTCCCGACTCTGTCTCTCTGGGAAATGGAATC	1260
QY	1261	TGGGAACCTGAAAGAGAGAGATTAACCTTGTGAAGAGCTGGGAGTGGCCAGTTTGA	1320
DB	1261	TGGGAACCTGAAAGAGAGAGATTAACCTTGTGAAGAGCTGGGAGTGGCCAGTTTGA	1320
QY	1321	GTGGTCCAGCTGGGCAAGTGAAGGGGCGAGTATGATGTTGCTGTTAGATGATCAAGAG	1380
DB	1321	GTGGTCCAGCTGGGCAAGTGAAGGGGCGAGTATGATGTTGCTGTTAGATGATCAAGAG	1380
QY	1381	GGCTCCATCTCAGAGATGAATTTCTTTCAGGAGGCCAGACTATGATGAACCTCAGCAT	1440
DB	1381	GGCTCCATCTCAGAGATGAATTTCTTTCAGGAGGCCAGACTATGATGAACCTCAGCAT	1440
QY	1441	CCCAAGCTGGTTAAATTTCTATGAGTGTGTTCAAAGAGATACCCCATATACATAGTACT	1500
DB	1441	CCCAAGCTGGTTAAATTTCTATGAGTGTGTTCAAAGAGATACCCCATATACATAGTACT	1500
QY	1501	GAATATATGAAGCAATGGCTGCTTGTGTAATACCTGAGGAGTCAAGGAAAGGACTTGA	1560
DB	1501	GAATATATGAAGCAATGGCTGCTTGTGTAATACCTGAGGAGTCAAGGAAAGGACTTGA	1560
QY	1561	CCCTCCAGCTCTTGAATAATGTCTACGATGCTGTGAAGGAGTGCCTTCTTGGAGAGT	1620
DB	1561	CCCTCCAGCTCTTGAATAATGTCTACGATGCTGTGAAGGAGTGCCTTCTTGGAGAGT	1620
QY	1621	CACCAATTCATACACCGGACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1680
DB	1621	CACCAATTCATACACCGGACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1680
QY	1681	GTGAAGATCTGACTTTTGAATGACAAAGTATGTTCTTGTATGACAGTATGTCAGTTCA	1740
DB	1681	GTGAAGATCTGACTTTTGAATGACAAAGTATGTTCTTGTATGACAGTATGTCAGTTCA	1740
QY	1741	GTGGAACAAAGTTTCCAGTCAAGTGGTCCAGCTCCAGAGTGTTCATTTCAATAC	1800
DB	1741	GTGGAACAAAGTTTCCAGTCAAGTGGTCCAGCTCCAGAGTGTTCATTTCAATAC	1800



1801 AGCAGCAAGTCAGACGATGGGATTTGGGATCCTGATGTGGAGGTGTTTCAGCCTGGG 1860  
 1801 AGCAGCAAGTCAGACGATGGGATTTGGGATCCTGATGTGGAGGTGTTTCAGCCTGGG 1860  
 1861 AAGCAGCCCTATGACCTTGTATGACAACTCCAGGTGTTCTGAAAGGTCTCCAGGGCCAC 1920  
 1861 AAGCAGCCCTATGACCTTGTATGACAACTCCAGGTGTTCTGAAAGGTCTCCAGGGCCAC 1920  
 1921 AGGCTTTACCGCCCGCCAGCTGGGATCGGACACCACTCTACGATCATGTACAGCTGCTGG 1980  
 1921 AGGCTTTACCGCCCGCCAGCTGGGATCGGACACCACTCTACGATCATGTACAGCTGCTGG 1980  
 1981 CAGCAGCTTCCAGAAAGCGTCCACATTTTTCAGCAACTCTGCTCTCCATTTGAACCACTT 2040  
 1981 CAGCAGCTTCCAGAAAGCGTCCACATTTTTCAGCAACTCTGCTCTCCATTTGAACCACTT 2040  
 2041 CGGGAAAAGACAGCAATTTGAAGAAGAAATTTAGGAGTCTGATGAAGATGAATATAGATG 2100  
 2041 CGGGAAAAGACAGCAATTTGAAGAAGAAATTTAGGAGTCTGATGAAGATGAATATAGATG 2100  
 2101 CTGGCAGCAGCTTTCATTTTATGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2160  
 2101 CTGGCAGCAGCTTTCATTTTATGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2160  
 2161 TTTTAAATAGTGTCTCTGATTTGCTATTTTGAAGATGAAGTAAAGTAAAGTAAAGTAAAG 2220  
 2161 TTTTAAATAGTGTCTCTGATTTGCTATTTTGAAGATGAAGTAAAGTAAAGTAAAGTAAAG 2220  
 2221 ATTCCTTTGAAATTTAGATCAATTTAGTAAATTTTGTGCTGCTGCTGCTGCTGCTGCTG 2280  
 2221 ATTCCTTTGAAATTTAGATCAATTTAGTAAATTTTGTGCTGCTGCTGCTGCTGCTGCTG 2279  
 2281 TTTCCAGCCTATAGCAGACACATTTTTCAGACTGCAATATAGAGACTGCTGCTGCTGCTG 2340  
 2281 TTTCCAGCCTATAGCAGACACATTTTTCAGACTGCAATATAGAGACTGCTGCTGCTGCTG 2340  
 2341 AAAGACTGAGCAGAACTGAAATAATTTTACTTATTTGATATTTTCTTCTTTTATATGTC 2400  
 2340 AAAGACTGAGCAGAACTGAAATAATTTTACTTATTTGATATTTTCTTCTTTTATATGTC 2399  
 2401 ATTGTCAACAATTTAATATATCTACCAAGTACAGAAATGTGAAA 2446  
 2400 ATTGTCAACAATTTAATATATCTACCAAGTACAGAAATGTGAAA 2445

RESULT 5

AAT00617  
 ID AAT00617 standard; cdna; 2494 BP.  
 AC AAT00617;  
 XX XX  
 DT 26-MAR-1996 (first entry)  
 XX XX  
 DE Megakaryocyte kinase MKK2 cdna.  
 XX XX  
 KW Megakaryocyte kinase-2; MKK2; cytoplasmic tyrosine kinase;  
 KW cellular signal transduction; leukaemia; myelofibrosis;  
 KW gene therapy; ss.  
 XX XX  
 OS Homo sapiens.  
 XX XX  
 FH Key Location/Qualifiers  
 DE CDS 76..2103  
 FT /\*tag= a  
 XX XX  
 PN W09529185-A1.  
 XX XX  
 PD 02-NOV-1995.  
 XX XX  
 PF 24-APR-1995; 95WO-US05008.  
 XX XX  
 PR 21-APR-1995; 95US-0426509.

22-APR-1994; 94US-0232545.  
 (PIAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 (SUGE-) SUGEN INC.  
 Gishizky M, Sures I, Ullrich A;  
 WPI: 1995-382959/49.  
 P-PSDB; AAR84182.  
 New poly:nucleotide(s) encoding megakaryocyte tyrosine kinase(s) -  
 used to develop prods. for the treatment and diagnosis of kinase  
 related signal transduction abnormalities.  
 Claim 4; Fig 2A-B; 82pp; English.  
 Overlapping cDNA clones were combined to obtain a sequence (AAT00617)  
 - coding for human megakaryocyte kinase MKK2 (AAR84182). The clones  
 were isolated from a fetal brain library using degenerate primers  
 CC (AAT00614-15) based on conserved regions within the kinase domain of  
 CC receptor tyrosine kinases. MKK polynucleotides can be used in the  
 CC prodn. in host cells of recombinant MKK, and in the gene therapy of  
 CC diseases such as acute megakaryocytic leukaemia, myelofibrosis and  
 CC acute megakaryocytic myelosis.  
 XX  
 SQ Sequence 2494 BP; 797 A; 504 C; 553 G; 640 T; 0 other;  
 Query Match 97.3%; Score 2388.6; DB 16; Length 2494;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 2405; Conservative 0; Mismatches 9; Indels 1; Gaps 1;  
 QY 19 GAGCAGCGATGATATATGATACAAAATCTATTCTAGAGAACTCTTCTCTCAAAAGATCA 78  
 DB 61 GAGTAGGATGATATATGATACAAAATCTATTCTAGAGAACTCTTCTCTCAAAAGATCA 120  
 QY 79 CAGCAAAAGAGAAATGTCACCAATTAATACAAAGAACGGCTTTTGTGTTGACCAAA 138  
 DB 121 CAGCAAAAGAGAAATGTCACCAATTAATACAAAGAACGGCTTTTGTGTTGACCAAA 180  
 QY 139 ACAACCTTTCTACTACTGATATGATACAAAATGAAAGGGGAGGAGGAGGAGGAGGAGG 198  
 DB 181 ACAACCTTTCTACTACTGATATGATACAAAATGAAAGGGGAGGAGGAGGAGGAGGAGG 240  
 QY 199 GAAATTAAGAAATCAGATGTGTGAGAAAGTAAATCTCGAGAGCAGACGCTGTAGAG 258  
 DB 241 GAAATTAAGAAATCAGATGTGTGAGAAAGTAAATCTCGAGAGCAGACGCTGTAGAG 300  
 QY 259 AGACAGTACCCATTTTCAGATTTCTATAAAGATGGGCTTCTATGTCTATGCATCAAT 318  
 DB 301 AGACAGTACCCATTTTCAGATTTCTATAAAGATGGGCTTCTATGTCTATGCATCAAT 360  
 QY 319 GAAGAGAGCCGAGTACAGTGTGAAAGCATTTACAAAAGAGATAGGGGTAAACCCAC 378  
 DB 361 GAAGAGAGCCGAGTACAGTGTGAAAGCATTTACAAAAGAGATAGGGGTAAACCCAC 420  
 QY 379 CTGCTGGTCAAGTACCATAGTGGGTTCTTCGTGGAGCGGAACTTCTGTGTTGCCAGCAG 438  
 DB 421 CTGCTGGTCAAGTACCATAGTGGGTTCTTCGTGGAGCGGAACTTCTGTGTTGCCAGCAG 480  
 QY 439 AGCTGTAAAGCAGCCCCCAGGATGTACCTCTCTGGGAGCATATGCTAAATCTGCATATGCA 498  
 DB 481 AGCTGTAAAGCAGCCCCCAGGATGTACCTCTCTGGGAGCATATGCTAAATCTGCATATGCA 540  
 QY 499 GTCATGAGAGAAACACAGAGTTCCTCCACCTTCCAGACAGAGTCTGATAGATACCTCGG 558  
 DB 541 GTCATGAGAGAAACACAGAGTTCCTCCACCTTCCAGACAGAGTCTGATAGATACCTCGG 600  
 QY 559 GCAGTCTCTGTCTCAAAATGGATGCACCATCTTCAAGTACCACCTCTAGCCCAATATGAC 618  
 DB 601 GCAGTCTCTGTCTCAAAATGGATGCACCATCTTCAAGTACCACCTCTAGCCCAATATGAC 660  
 QY 619 AACGAATCAAGAAAAAATATGCTGCTCCACAGCCACCATCTTCAAGTACCACCTCTAGCCCA 678



XX	22-OCT-2001; 2001WO-US51415.
PF	
XX	
XX	23-OCT-2000; 2000US-242471P.
XX	
PA	(BRIM ) BRISTOL-MYERS SQUIBB CO.
XX	
PI	McAtee P;
XX	
PS	WPI; 2002-519317/55.
DR	P-PSDB; ABB76488.
XX	
PT	Assay for identifying compound that modulates the activity of Bruton's
PT	tyrosine kinase, by contacting the cell expressing kinase with a test
PT	compound, and determining whether the compound modulates the kinase
PT	activity
XX	
XX	Disclosure; Page 13-14; 60pp; English.
PS	
XX	The present sequence is the nucleotide sequence of the murine
CC	Bruton's tyrosine kinase (BTK) gene encoding a 539-amino acid
CC	polypeptide (see ABB76488). BTK and intermediates in the BTK
CC	pathway are critical intermediates in the cytoskeletal
CC	rearrangement pathway leading to osteoclast activation. Mice
CC	deficient in BTK exhibit osteopenia, which can be reversed upon
CC	addition of multiple copies of the BTK gene in transgenic mice.
CC	Modulators of BTK activity and BTK intermediate activity are
CC	therefore useful in affecting osteoclast activity and bone
CC	resorption. Such modulators may be identified using assays of the
CC	present invention, and are expected to be useful in the treatment
CC	of osteoporosis and related disease states. BTK target validation
CC	studies on modulators may be carried out using conventional
CC	osteoporosis mouse models. The invention also includes the use of
CC	antisense therapy.
XX	
SQ	Sequence 2468 BP; 713 A; 559 C; 602 G; 594 T; 0 other;
	Query Match            20.1%; Score 494.6; DB 24; Length 2468;
	Best Local Similarity   64.0%; Pred No. 3.8e-125;
	Matches       746; Conservative   0; Mismatches   419; Indels   0; Gaps   0;
QY	887 CTGAAGAAGAGGAAAACCTGGATGATTACACTGTGTTTGTGTGTAACTTCAGATC
Db	941 CTGAAGCTGAGGACTCCATAGAGATGTATCAGTGGTATTCCAAGCATGACTCGAAGTC
QY	947 AATCTGAACAGTTACTCAGCAACAAGGGAAAAAGAGGAGGCATTATTGGTTAGAAATCGA
Db	1001 AAGCTGAGCAACTGCTAAAGCAGAGGGGAAAGAGGAGGTTTCATTGTGCAGAGACFCCA
QY	1007 GCCAAGTGCGAAATGTACACAGTGTCTCTATTATTAGTAAGGCTGTGAATGATAAAAAAGGAA
Db	1061 GCANAAGCTGGAATAATACACCCTGCTGTGTTTGTCTTAATCTACTGGGAGCCTCAAGGGG
QY	1067 CTGCTCAAACTATACCAGTGCATCAAAAGCTGAGAACAAATATACCTGGCAGAAACT
Db	1121 TGATCGGCCAATTACGTTGTGTGTTCCACGCCACAGAGCCAGTAGTATTACCTGGCTGAGAAAC
QY	1127 ACTGTTTTGATTCCATTCCAAAGCTTATTCATTATCATCAACACAATTACAGCAGGCATGA
Db	1181 ACCTCTTACGACGACATCCCCTGNGCTCATTTAACTACCATCAACACAACTCTGCAGGCTCA
QY	1187 TCACAGGCTCCGCCACCCTGTGTCAACAAGGCCAACAGGTCGCCGACTCTGTGTGCC
Db	1241 TATCCAGGCTGAATAATATCCTGTGTCYNAACAAACAAACAGCGCCCTTCTACTGCAGGCC
QY	1247 TGGGAAATGGAAATCTGGGAACTGAAAAGAGAAAGAGATTACCTGTTTGAAGGAGCTGGGAA
Db	1301 TGGGCTATGTCATCGSAAATTTGATCCAAAGGACCTCACCTTCTTGAAGAGGCTTGGGA
QY	1307 GTGGCGAATTTGGAGTGGTCCAGCTGGGCAAGTGGAGGGGCGAGTAGTAGTGTGCTGTTA
Db	1361 CTGGACAAATTCGGTGTGCTGAAATATGGAAGTGGAGGGGCCAATATGATGTGGCCATCA

FT	polyA_signal	/*tag- c	2458..2463	
FT	polyA_site	/*tag- d	2485	
FT		/*tag- e		
XX	US5550054-A.			
PN	27-AUG-1996.			
XX	21-FEB-1995;	95US-0391615.		
XX	21-JAN-1993;	93US-0006449.		
PR	04-DEC-1992;	92US-0985998.		
PR	21-FEB-1995;	95US-0391615.		
XX	(REGC ) UNIV CALIFORNIA.			
PA	Rawlings D, Saffran D, Tsukada S, Witte O;			
PI	WPI; 1996-401601/40.			
XX	P-PSDB; AAW06708.			
DR	New haematopoietic restricted tyrosine kinase DNA - used to develop			
PT	products for modulating tyrosine kinase activity, e.g. for treating			
PT	neoplasia, auto-immune disease or allergy			
XX	Claim 2; Column 19-24; 26pp; English.			
PS	AA145648 is a cDNA sequence encoding a murine haematopoietic-			
XX	restricted tyrosine kinase (TK) also known as B cell progenitor			
CC	kinase (BPK). The TK enzyme can be used to identify compounds that			
CC	inhibit or stimulate the expression and/or function of TK. The			
CC	TK is capable of autophosphorylation and transphosphorylation.			
CC	Modulators of TK activity can be used to treat neoplasia and			
CC	non-neoplastic hyperproliferative disease, e.g. allergy and			
CC	autoimmune disease. TK may also be used to determine cell-types			
CC	based on the nature and amount of TK present.			
CC	(Updated on 25-MAR-2003 to correct PF field.)			
XX	Sequence 2505 BP; 743 A; 559 C; 602 G; 601 T; 0 other;			
QQ	Query Match	20.1%;	Score 493;	DB 17; Length 2505;
	Best Local Similarity	63.9%;	Pred No. 1.1e-124;	
	Matches 745; Conservative 0; Mismatches 420; Indels 0; Gaps 0;			
QY	887 CTGAAGAGAGGAAACCTGGATGATTATGACTGTTTGGTGAACATCTCCAGATCAC	946		
DB	945 CTGAAGCTGAGGACTCCATAGAGATGATGAGTGTATCCCAAGCACATGACTCGAAGTC	1004		
QY	947 AATCTGAACAGTTACTCAGACAAAGGAAAGAGAGGACATTTATGGTTAGAAATTCGA	1006		
DB	1005 AAGCTGAGCAACTGCTAAGCAAGAGGAAAGAGAGGTTTCATTGTCAGAGACTCCA	1064		
QY	1007 GCCAAGTGGGAATGPACACAGTGTCTTATTAGTAAAGCTGTGATGATAAAAGGNA	1066		
DB	1065 GCAAGCTGGAAATATACACCGTGTCTGTGTTGCTAAATCTACTGGGGAGCCTCAAGGG	1124		
QY	1067 CTGTCAACATATACACAGTGTGATCAAAATGCTGGAACAAATATATACCTGGCAGAAACT	1126		
DB	1125 TGATCCGCCATTACCTGTGTGTTCCAGCCACAGAGCAGTATTACCTGGCTGAGAAAC	1184		
QY	1127 ACTGTTTGTATCCATTCCAAGCTTATTCATTATCATCAACAAATTCAGAGCATGA	1186		
DB	1185 ACCTCTTCAGCACCACCTCCTGAGCTCAATTAATACATCAACAACTCTGAGCCCTCA	1244		
QY	1187 TCACAGGCTCCGACACCTGTGTCAACAAAGGCAAGTCCCGACTCTGTGTCC	1246		
DB	1245 TATCAGGCTGAATATATCTGTCTTAACAAACAAACGCGCTCTCTGTCAGGCGC	1304		
QY	1247 TGGGAATGGAATCTGGCACTGAAAGAGAGAGATTTACCTTGTGAGGAGCTGGNA	1306		
DB	1305 TGGCTATGGATATGGGAAATTTGATCCAAAGGACCTCACCTTCTTGAAGGAGCTTGGGA	1364		

QY	1307 GTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGGAGGGCAGTATGATGTTGCTGTTA	1366
DB	1365 CTGGACAATTCGGTGTCTGTAATATGGAAGTGGAGGGCCAAATATGATGTGCCATCA	1424
QY	1367 AGATGATCAAGAGAGGCTCCATGTCAGAAATGAATTTCTTTCAGGAGGCCAGACTATGA	1426
DB	1425 AGATGATCAGAGAAGGTTCCATCTCGAGGATGAATTCATTGAGAGAAGCCAAAGTCATGA	1484
QY	1427 TGAACACTCAGCCATCCCAAGCTGTTAAATTTCTATGAGGTGTGTTCAAGGAATACCCCA	1486
DB	1485 TGAATCTTTCCCATGAGAAGCTGTCGAGTGTGATGGCTGTGCACCAACAAAGCCGCCA	1544
QY	1487 TATACATAGTACTGAATATATATAAGCAATGGCTTGTCTGTAATTAACCTGAGGAGTCA	1546
DB	1545 TCTTCATCATCACCGAGTACATGGCTAAATGGCTTCTTGAATCTACCTGAGGAGATGC	1604
QY	1547 GAAAGGACTTGAACCTTCCAGCTCTTAGAATGTGCTAGCATGTCTGTGAAGGCATGG	1606
DB	1605 GGCACCGCTTCCAGACACAGCAGCTGCTTGAGATGTGCAAGATGTCTGTCAAGCAATGG	1664
QY	1607 CCTTCTTGGAGAGTCCACCAATTCATACACCGGACTTGGCTGCTGTAACCTGCTTGGTGG	1666
DB	1665 AATACTTGGAGTCGAGAGCTTCTTCAACAGACCTGGAGCTCGAAGACTGTTGGTAA	1724
QY	1667 ACAGAGATCTCTGTGTGAAAGTATCTGACPTTGGAAATGACAAGGTATGTTCTTGATGACC	1726
DB	1725 ACGATCAAGGAGTTGTGAAAGTATCTGACPTTGGCTGCTAGGTATGTCTTCTGATGATG	1784
QY	1727 AGTATGTCAGTTCAGTCGAGCAAAAGTTTCCAGTCAAGTGGTTCAGCTTCAGAGGTGTTTC	1786
DB	1785 AGTACACCGCTCTGTAGGCTCCAAAGTTTCCAGTTCGGTGGTCTCCACAGAGAGTGTCTTA	1844
QY	1787 ATTACTTCAATACACAGCAGCAAGTCAAGCTATGGGCTATGGGATCTGATGTGGAGG	1846
DB	1845 TGTATAGCAAGTTCACAGCAATCTGACATCTGGGCTTTTGGGTTTAAATGTGGAGAGA	1904
QY	1847 TGTTCAGCTGGGGAAGCAGCCCTATGACPTTGTATGACAACTCCCAAGTGGTCTGAAGG	1906
DB	1905 TCTACTCCCTGGGGAAGATCCGCTATGAGAGATTTACTAACAGTACAGCAGCAACACA	1964
QY	1907 TCTCCAGGCGCACAGGCTTACCGGCCACCTGGCATCGACACCATCTACACAGATCA	1966
DB	1965 TTGCTCAAGCTTACCTCTCTACAGGCTCTCTGATCAGAGAGGTTATATACCATCA	2024
QY	1967 TGTACAGCTGCTGGCAGCAGCTTCCAGAAAAGCGTCCACATTTTCAGCAACTCTGCTTT	2026
DB	2025 TGTACAGCTGCTGGCAGCAGAGAGCAGATCAACGCTCCTAGTTTCAAAATTTCTCTTGAGTA	2084
QY	2027 CCATTGAACCACTTCGGGAAAGAGA	2051
DB	2085 ACATTCTAGATGTGATGGATGAAGA	2109

RESULT 8

ABK85707

ID ABK85707 standard; DNA; 1365 BP.

XX ABK85707;

XX ABK85707;

DT 20-AUG-2002 (first entry)

XX DNA encoding truncated Tec Kinase Btk.

XX Tec kinase; pleckstrin-homology domain; Tec homology domain;  
 XX inflammation; asthma; allergic rhinitis; arthritis; rheumatoid arthritis;  
 KW adult respiratory distress syndrome; osteoarthritis; uveitis;  
 KW inflammatory eye condition; inflammatory bowel condition;  
 KW Crohn's disease; ulcerative colitis; periodontal disease; oesophagitis;  
 KW inflammatory skin condition; psoriasis; eczema; dermatitis; Btk;  
 gene; ds; human.  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 1..1365  
 FT /\*tag= a  
 FT /product= "Truncated Btk"  
 FT /note= "tec kinase"  
 XX WO200234899-A2.  
 XX  
 PD 02-MAY-2002.  
 XX  
 XX 17-OCT-2001; 2001WO-EPI1949.  
 XX  
 XX 20-OCT-2000; 2000GB-0025804.  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX  
 XX Sims MJ, Hayes DJ;  
 XX  
 XX WPI; 2002-471439/50.  
 DR P-PSDB; AAU98426.  
 XX  
 PT New truncated Tec kinase polypeptide useful for identification of a  
 PT compound which modulates the activity of the polypeptide for treating  
 PT inflammatory conditions such as asthma, rheumatoid arthritis, and  
 PT uveitis  
 XX  
 XX Claim 13; Fig 9; 38pp; English.  
 XX  
 CC The invention describes a truncated Tec kinase polypeptide (I) having a  
 CC Tec kinase amino acid sequence truncated by a minimum of the amino acids  
 CC constituting the N-terminal pleckstrin-homology (PH) domain and a portion  
 CC of the Tec homology (TH) domain including a proline rich region up to but  
 CC not including the amino acids constituting the kinase domain. A compound  
 CC (II) which modulates Tec activity is useful in treatment of a disorder  
 CC such as inflammation that is responsive to modulation of Tec  
 CC kinase activity. (I) is useful for screening for targets for the  
 CC identification and development of new pharmaceutical agents, including  
 CC agonists and antagonists of Tec kinases, which may be useful in therapy.  
 CC The nucleic acid (II) encoding (I) is useful for treating inflammatory  
 CC diseases such as asthma, allergic rhinitis, adult respiratory distress  
 CC syndrome, arthritic conditions such as rheumatoid arthritis,  
 CC osteoarthritis, inflammatory eye conditions such as uveitis, inflammatory  
 CC bowel conditions such as Crohn's disease, ulcerative colitis, periodontal  
 CC disease, oesophagitis, and inflammatory skin conditions such as  
 CC psoriasis, eczema and dermatitis. (I) is suitable for screening methods  
 CC as they do not need to be pre-activated by phosphorylation. In vivo, Tec  
 CC kinases need to be phosphorylated by other kinases in order to activate  
 CC the enzyme, while pre-activation by phosphorylation is commonly required  
 CC in assays. The removal of the need to pre-activate (I) offers a  
 CC simplification for the assay. A further advantage of (I) is the provision  
 CC of an assay that is robust as: it is possible to generate large amounts  
 CC of truncated enzyme which are stable over a long time and the assay gives  
 CC a high frequency of comparable results upon repeat testing. This sequence  
 CC encodes a truncated construct of the Tec kinase Btk.  
 XX  
 SQ Sequence 1365 BP; 395 A; 301 C; 344 G; 325 T; 0 other;  
 Query Match 19.78; Score 485; DB 24; Length 1365;  
 Best Local Similarity 63.5%; Pred. No. 1.2e-122;  
 Matches 740; Conservative 0; Mismatches 425; Indels 0; Gaps 0;  
 QY 887 CTGAAGAAGAGGAAACCTGGATGATTGACTGGTTTGGTGGTAAACATCTCCAGATCAC 946  
 DB 194 CTGACGGGAGACTCCATAGAAATGATGAGTGGTATTCCAAACACATGACTCGGAGTC 253  
 QY 947 AATCTGACACAGTTACTCAGCAAGGAAAGGAGGAGGATTTATGGTTAGAAATTCGA 1006  
 DB 254 AGGCTGAGCAACTGCTTAAGCAAGAGGGGAAAGAGGAGGTTTCATTGTGACGACTCCA 313  
 QY 1007 GCCAAGTGGGAATGTACACAGTGTCTTATTATTAGTAGGCTGTGAATGATAAAAGGAA 1066  
 DB 314 GCAAAGCTGGCAAAATATACAGTGTCTGTGTTTGTGTAATCCACAGGGGACCCTCAAGGGG 373

RESULT 9  
 ABN83857  
 ID ABN83857 standard; DNA; 2560 BP.

QY 1067 CTGTCAACAATACACAGTGCATCAAAATGCTGAGAACAAATATACCTGGCAGAAAACCT 1126  
 DB 374 TGATACGTCATATATGTTGTGTTCACACACCTCAGAGCCAGTATTACCTGGTGAGAAGC 433  
 QY 1127 ACTGTTTGTATCCATCCCAAGCTTATTCAATCAACACAAATTCAGCAGGCATCA 1186  
 DB 434 ACCTTTTCAGCACCATCCCTGAGCTCATTAATACCATCAGCACAACCTCTGCAGGACTCA 493  
 QY 1187 TCACAGGCTCCGCCACCTGTGTCAACAAGAGGCCAACAAAGTCCCGACTCTGTGTCC 1246  
 DB 494 TATCCAGGCTCAATATCCAGTGTCTCAACAACAAGATGACCTTCCACTGCAGGCC 553  
 QY 1247 TGGAAATGGAATCTGGGAACCTGAAAGAGAGATTAACCTTGTGAAGAGCTGGGAA 1306  
 DB 554 TGGGATACGGATCATGGGAATTGATCCAAAGGACCTGACCTTCTTGAAGAGCTGGGA 613  
 QY 1307 GTGCCAGTTTGGAGTGTCCAGCTGGCAAGTGGAAAGGCCAGTATGATTTGCTGTTA 1366  
 DB 614 CTGGACAATTTGGGTAGTGAAGTATGGGAATGGAGAGGCCAGTACGACGTGGCCATCA 673  
 QY 1367 AGATGATCAAGGAGGCTCCATGTCAAGAGATGAATTTCTTCAGGAGGCCACAGATATGA 1426  
 DB 674 AGATGATCAAGAGGCTCCATGTCTGAAGATGAATTCATTGAAGACCAAGATCATGA 733  
 QY 1427 TGAACATCAGCCATCCCAAGCTGGTTAAATTTCTATFGAGTGTGTTCAAAGGAATACCCCA 1486  
 DB 734 TGAATCTTCCCATGAGAAAGCTGGTGACAGTTGTATGGGCTCTGCACCAAGCAGCGCCCA 793  
 QY 1487 TATACATAGTACACTGAATATATAGCAATGGCTGCTGTGTAATTAACCTGAGGAGTCAAG 1546  
 DB 794 TCTTCATCATCAGTACATGGCAATGGCTGCTCTTCACTACCTGAGGAGAGATC 853  
 QY 1547 GAAAGGACTTGAACCTTCCAGCTCTTAGAAATGTCTACGATGTCTGTGAAGGATGG 1606  
 DB 854 GCCACCGCTTCAGACTCAGCAGCTGTAGAGATGTGCAAGATGTCTGTGAAGCCATGG 913  
 QY 1607 CTTTCTTGGAGAGTACCAATTCATACACCGGACTTGGCTGCTCGTAACTCCTTGGTGG 1666  
 DB 914 AATACCTGGAGTCAAGCAGTTCCTTCACCGAGACCTGCGAGCTCGAAACTGTTGGTAA 973  
 QY 1667 ACAGAGATCTCTGTGTGAAGTATCTGACTTTGGAAATGACAAGGATGTGTTGTGAGACC 1726  
 DB 974 ACGATCAAGGAGTGTGTTAAAGTATCTGATTTGGGCTGTCCAGGATGTCTCTGGATGATG 1033  
 QY 1727 AGTATGTCAGTTCAGTCGGACAAAGTTTCCAGTCAAGTGGTCACTCCAGAGGTCTTTC 1786  
 DB 1034 AATACAAAGTCTAGTAGGCTCCAAATTTCCAGTCCGGTGGTCCCGGAGGTCTCTGA 1093  
 QY 1787 ATTACTTCAAAATACAGACAGCAAGTCAGACGTATGGGCATTTTGGGATCTCTGATGGGAG 1846  
 DB 1094 TGTATAGCAAGTTCAGCAGCAAACTGACATTTGGGCTTTTGGGGTTTGTATGTGGGAAA 1153  
 QY 1847 TGTTCAGGCTGGGGAAGCAGCCCTATGACTTGTATGACAACCTCCAGGTTGTTCTGAAGG 1906  
 DB 1154 TTTACTCCCTGGGGAAGATGCCATATAGAGATTTTACTAAACAGTGGACTGTGTGAACACA 1213  
 QY 1907 TCTCCAGGGCCACAGGCTTTACCGGCCCTCCAGTCCGACACCACTCTACCATCA 1966  
 DB 1214 TTGCCCCAAGGCTTACGCTCTACAGGCTCTCATCTGGCTTCAGAGAGGTATATACCATCA 1273  
 QY 1967 TGTACAGTGTGGCAGCAGCTTCCAGAAAAGGCTCCACCATTTTCAGCAACTCTCTCTT 2026  
 DB 1274 TGTACAGTGTGGCATGAGAAAGCAGATGAGCGTCCCATCTTCAAAATTCCTCTCTGAGCA 1333  
 QY 2027 CCATTGAACCACTTCGGGAAAAAGA 2051  
 DB 1334 ATATTCTAGATGTCATGGATGAAGA 1358





AC AAS14047;  
 XX 18-DEC-2001 (first entry)  
 XX Human cDNA encoding Bruton's tyrosine kinase, Btk.  
 DE  
 XX Immunosuppressive; antiinflammatory; osteopathic; antirheumatic;  
 KW antiarthritic; dermatological; antipsoriatic; antiasthmatic; cytostatic;  
 KW cardiant; antiarteriosclerotic; vasotropic; cerebroprotective;  
 KW neuroprotective; nontropic; ophthalmological; immunosuppressive;  
 KW hepatotropic; virucide; anti-HIV; immunomodulator; antithyroid; sepsis;  
 KW septic shock; inflammation; Crohn's disease; rheumatoid arthritis;  
 KW leukopaemia; osteoarthritis; spondyloarthropathy; psoriatic arthritis;  
 KW severe steroid resistant asthma; pulmonary fibrosis; myocarditis;  
 KW atherosclerosis; angina; vasculitis; reperfusion injury; infarction;  
 KW stroke; multiple sclerosis; Alzheimer's disease; Graves ophthalmopathy;  
 KW psoriasis; transplant rejection; acute alcoholic hepatitis; fibrosis;  
 KW HIV; human immunodeficiency virus; cachexia; erythema nodosum lepromatous;  
 KW borreliosis; meningococcal septicaemia; cancer; Human;  
 KW Bruton's tyrosine kinase; Btk; ss; antibacterial.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 164..2071  
 FT /\*tag= a  
 FT /product= "Btk"  
 FT  
 XX W0200166107-A2.  
 XX  
 XX 13-SEP-2001.  
 XX  
 XX 06-MAR-2001; 2001WO-GB00949.  
 XX  
 XX 06-MAR-2000; 2000GB-0005345.  
 XX  
 XX (KENN-) KENNEDY INST RHEUMATOLOGY MATHILDA & TER.  
 XX  
 XX Foxwell BMJ;  
 XX  
 XX WPI; 2001-602596/68.  
 DR P-PSDB; AAU09009.  
 XX  
 XX Treating condition associated with cytokine production in a mammal,  
 PT e.g. sepsis, septic shock, inflammation and Crohn's disease, comprises  
 PT administering an inhibitor of a member(s) of the Tec family of protein  
 PT tyrosine kinases  
 XX  
 XX Claim 23; Fig 11; 78pp; English.  
 PS  
 XX The invention relates to the use of inhibitors of the Tec family of  
 CC protein tyrosine kinases to treat a condition associated with cytokine  
 CC production in a mammal. The Tec family members are e.g. Bruton's  
 CC tyrosine kinase (Btk), Tec, Itk, Tsk or Bmx. The condition is especially  
 CC sepsis, septic shock, inflammation, Crohn's disease or a condition  
 CC induced by TRR (Toll related receptor) ligand, LPS (lipopolysaccharide),  
 CC zymosan or gram-negative bacteria. The method may also be used for  
 CC treating rheumatoid arthritis. Biological effects of LPS include fever,  
 CC circulatory disturbances and vascular hypersensitivity to adrenergic  
 CC drugs. Leukopenia typically followed by leucocytosis, non-specific  
 CC stimulation of B-lymphocytes to undergo blast transformation and  
 CC proliferation, lethal toxicity and non-specific tolerance to endotoxin  
 CC through repeated exposure to LPS. LPS also causes an increase in  
 CC production of cytokines e.g. TNF. Conditions associated with IL-1beta  
 CC and TNFalpha osteoarthritis, spondyloarthropathy, psoriatic arthritis,  
 CC severe steroid resistant asthma, pulmonary fibrosis, myocarditis,  
 CC atherosclerosis, angina with high CRP/IL-6, vasculitis, reperfusion  
 CC injury, infarction, stroke, multiple sclerosis, Alzheimer's disease,  
 CC Graves ophthalmopathy, psoriasis, transplant rejection, acute alcoholic  
 CC hepatitis, fibrosis. Other conditions for which inhibitors of Tec  
 CC family PKs may be used include HIV (human immunodeficiency virus  
 CC infection), cachexia, erythema nodosum lepromatous, borreliosis, or  
 CC meningococcal septicaemia, cancer, such as breast cancer, ovarian cancer

CC or colon cancer. Many other examples of diseases are given in the  
 CC specification. The present sequence encodes human Btk.  
 XX  
 SO Sequence 2582 BP; 750 A; 573 C; 632 G; 627 T; 0 other;  
 Query Match 19.6%; Score 480.2; DB 22; Length 2582;  
 Best Local Similarity 63.3%; Pred. No. 3.6e-121;  
 Matches 737; Conservative 0; Mismatches 428; Indels 0; Gaps 0;  
 QY 887 CTGAAGAGAGAGAAACCTGGATGATTTATGACTGGTTTCTGGTACATCTCCAGATCAC 946  
 DB 972 CTGAAGAGAGAGACTCCATAGAAATGTATGATGGTATTTCAACACACATGACTCGGAGTC 1031  
 QY 947 AATCTGAACAGTTACTCAGACAAAGGAGAGAGGAGCATTTATGGTTAGAAATTCGA 1006  
 DB 1032 AGGCTGACCACTCTAAGCAAGAGGGAAGAGAGGTTTCATTGTCAGAGACTCA 1091  
 QY 1007 GCCAAGTGGGAATGTACACAGCTGCTCTTATTAGTAAAGGCTGTGAATGATAAAAGGAA 1066  
 DB 1092 GCAAAGCTGGCAAAATATACAGTGTCTGTGTTGCTTAAATCCACAGGGGACCTCAAGGGG 1151  
 QY 1067 CTGTCAACATTTACCAGTGCATACAAATGCTGACAAACAAATTTATACCTGGCAGAAACT 1126  
 DB 1152 TGATACGTCATTTATGTTGTTCCACACCTCAGAGCCAGTATTACCTGGCTGAGAAGC 1211  
 QY 1127 ACTGTTTGTATTCATTCCTCAAGGCTTATTCATTTATCATCAACACAAATTCAGCAGCATGA 1186  
 DB 1212 ACCTTTTCAGCACCATCCCTGAGCTCATTAATCACTCACCATCAGCACAACCTCTGAGACTCA 1271  
 QY 1187 TCACACGGCTCCGCCACCTGTGTCAACAAAGGCCAACAAAGTCCCGACTCTGTGTCC 1246  
 DB 1272 TATCCAGGCTCAAAATATCCAGTGTCTCAACAAACAAAGAAATGCACCTTCCACTGCGAGGCC 1331  
 QY 1247 TGGGAATGGAATCTGGGAATGAAAGAGAGAGATTTACCTTGTGTAAGGAGCTGGGAA 1306  
 DB 1332 TGGGATACGGATCATGGGAAATGTGATCCAAAGGACCTGACCTTCTTGAAGSAGCTGGGGA 1391  
 QY 1307 GTGCCAGTTTGGAGTGGTCCAGCTGGCAAGTGGAGGGGAGTATGATGTTGCTGTGA 1366  
 DB 1392 CTGACACATTTGGGAGTAGTGAAGTATGGGAATGAGAGAGCCAGTACGACGTGGCCATCA 1451  
 QY 1367 AGATGATCAAGAGGGCTCCATGTGCAAGATGAATTTCTTCAAGAGGCCAGACTATGA 1426  
 DB 1452 AGATGATCAAGAGGGCTCCATGTGCAAGATGAATTTCTTCAAGAGGCCAGACTATGA 1511  
 QY 1427 TGAACATCAGCCATCCCAAGCTGGTAAATTTCTATGAGTGTGTTTCAAGAGGAATACCCCA 1486  
 DB 1512 TGAATCTTCCCATGAGAGGCTGGTGCAGTTGTATGGCTGTGCAACAGCAGCGCCCA 1571  
 QY 1487 TATACATGACTGAATATATTAAGCAATGGCTGCTGCTGAATTAACCTGAGGAGTCAAG 1546  
 DB 1572 TCTTCATCATCATGAGTACATGGCCATGGCTGCTGCTGAACCTACCTGAGGAGATGC 1631  
 QY 1547 GAAAGAGCTTGAACCTTCCAGCTCTTTAGAAATGTGTACAGTGTGTGTGAAGGAGATGG 1606  
 DB 1632 GCCACCGCTTCCAGACTCAGCAGCTGTGAGATGTGCAAGGATGTCTGTGAAGGAGATGG 1691  
 QY 1607 CCTTCTTGGAGAGTACCAATTCATACACCGGAGCTGGCTGCTGTAACCTGCTTGGTGG 1666  
 DB 1692 AATACCTGGAGTCAAGAGCAGTTCTTCCAGGAGCTGGCAGCTCGAAACTGTTGGTAA 1751  
 QY 1667 ACAGAGATCTCTGTGTGAAAGTATCTGACTTTTGAATGACAAAGGTATGTTCTGTGATGACC 1726  
 DB 1752 ACGATCAAGAGAGTTGTTAAAGTATCTGATTTGGGCTGTCCAGGATGTCTCTGGATGATG 1811  
 QY 1727 AGTATGTCAAGTTCAGTTCGGAACAAAGTTTCCAGTCAAGTGGTCACTCCAGAGGTGTTTC 1786  
 DB 1812 AATACCAAGCTCAGTAGGCTCCAAATTTCCAGTCCGGTGGTCCCGCCAGGAGCTCTGA 1871  
 QY 1787 ATTACTTCAATACAGCAGCAGTATGAGAGTATGGGATTTGGGATTCCTGATGTGGGAGG 1846  
 DB 1872 TGTATAGCAAGTTTCAGCAGCAAAATCTGACATTTGGGCTTTTGGGGTTTGTGTGGGAA 1931





ID AAS70719 standard; cDNA; 3650 BP.  
AC AAS70719;  
XX  
XX 13-FEB-2002 (first entry)  
XX  
XX DNA encoding novel human diagnostic protein #6523.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
XX  
XX W0200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
PI  
XX  
XX WPI; 2001-639362/73.  
DR P-PSDB; ABG06532.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX Claim 1; SEQ ID No 6523; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcr\_sequences.  
XX  
XX Sequence 3650 BP; 1083 A; 726 C; 884 G; 957 T; 0 other;  
SQ

Query Match 16.08; Score 392.8; DB 23; Length 3650;  
Best Local Similarity 58.94; Pred. No. 4.8e-97;  
Matches 698; Conservative 0; Mismatches 482; Indels 6; Gaps 1;  
QY 899 AAAACCTGGATGATATGACCTGGTTTGTGCTAATCTCCAGATCACAATCTGAACAGT 958  
DB 836 ACACTTAGATCAATATGAATGGTATTCAGAAATATGAATAGAACGACGACGACAC 895  
QY 959 TACTCAGACAAAAGGAAAGAGGACGATTTATGGTTAGAAATTCGAGCCCAAGTGGGAA 1018  
DB 896 TCCTCCGACGTGAAGATAAAGAGGATGGTTTATGGTAAAGGATTCAGTCACACGCGT 955  
QY 1019 TGTACACAGTCTCTTTATTTAGTAAAGCTGTGTAATGATATAAAAGGAACACTCTCAACAT 1078

DB 956 TGTACACAGTCTCCCTTTATACCAAGTTGGAGGAGAGGTTTCATCGGCTTTTAGCGATT 1015  
QY 1079 ACCACGTGCA-----TACAAATGCTGAGAACAAATTTATACCTGGCAGAAAACTACTGTT 1132  
DB 1016 ATCATATAAAGGAAACAACAACATCTCCAAAGAAAGTATTACCTAGCTGAAAAACATGCTT 1075  
QY 1133 TTGATTCATTCCAAAGCTTATTTCATTATCATCAACACAAATTCAGCAGGCATGATCACAC 1192  
DB 1076 TTGGCTCCATTCTCTGAGATTATTGAATATCATTAAGCAACAATGCAGCAGGCTTGTCACCA 1135  
QY 1193 GGCTCCGACACCTCTGTCAACAAGGCGCAACAAGTCCCGACTCTGTGTCCCTGGGAA 1252  
DB 1136 GCTTCGGTACCCAGTTAGTGTGAAGGGAAGATGCACCCACTCTCAGGATTCAGCT 1195  
QY 1253 ATGGAATCTGGAACTGAAAGAGAGAGATTTACCTTTGTGAAGGAGCTGGGAAGTGGCC 1312  
DB 1196 ATGAGAAATGGCAGATTAAACCTTCAGAACTGACCTTTATGAGGGAATTTGGGAAGTGGAC 1255  
QY 1313 AGTTTGGAGTGGTCCAGCTGGCAAGTGGCAAGGCGGAGTATGATGTTGCTGTTAAGATGA 1372  
DB 1256 TGTTTGGAGTGGTGGAGGCTTGGCAATGCGGAGCCCGAGTACAAAGTCGCAATCAAAAGCTA 1315  
QY 1373 TCAGGAGGCTCCATGTCAAGAGATGAATTTCTTCAGGAGGCCAGACTATGATGAAC 1432  
DB 1316 TTCGGAAGTGCAATGTCCGAGGAGGACTTTATAGAAGAAGCTAAGTGTATGATGAAGC 1375  
QY 1433 TCAGCCATCCCAAGCTGGTTAAATTTCTATGGAGTGTGTTTCAAGGAATACCCCATATACA 1492  
DB 1376 TGACACACCCGAAGTTAGTGCAGCTTATGTGTGTGTCACCCAGCAGAAACCAATATACA 1435  
QY 1493 TAGTACTGAATATATAAGCAATGGCTGCTTCTGTAATTTACCTCAGGAGTACGCGGAAAG 1552  
DB 1436 TTGTTACTGAGTTTCATGGAAGGGCTGCTTCTGAAATTTCTCCACAGAGACAAGGTC 1495  
QY 1553 GACTTGAACCTTCCAGCTCTTAGAAATGCTGCTAGATCTCTGTGAAGGCATGCGCTTCT 1612  
DB 1496 ATTTTCAGTAGACAGTACTGCTGAGCATGTGTGAGGATGTGTGTAAGGGATGGAGTATC 1555  
QY 1613 TGGAGAGTCAACAAATTCATACACCGGAGTGGCTGCTGCTAACTGCTTGGTGACAGAG 1672  
DB 1556 TGGAGAGAAACAGCTTCATCCACAGATCTGGCTGCCAGAAATTTCTAGTAGAGGAGG 1615  
QY 1673 ATCTCTGTGTGAAGATCTGACTTTGGAAATGACAAGTATGTTTGTGATGACGAGTATG 1732  
DB 1616 CGGGAGTTGTAAGATATCTGATTTTGGAAATGCCAGGTATTTTCTGGATGATCAGTACA 1675  
QY 1733 TCAGTTTCAGTCGGAACAAGTTTCCAGTCAAGTGGTCAAGTCCAGAGTGTGTTTCATTA 1792  
DB 1676 CAAGTTCTTCTGGTGCTAAGTTTCTGTGAAGTGGTGTCCACTGAAAGTGTGTTTAATACA 1735  
QY 1793 TCAATATACAGCAGCAAGTCAGACGATGGGCATTTGGGATCTGTGATGTGGAGGTGTCA 1852  
DB 1736 GCGCTTTCAGCAGCAATCAGATGCTGCTCATTTGGTGTGTTTAAAGTGGAGTATCA 1795  
QY 1853 GCCTGGGAAGCAGCCCTATGACTTGTATGACAACCTCCAGGTGTGTTTGAAGGTCTCCC 1912  
DB 1796 CGGAAGGCAGAAATGCGCTTTTGAAAAATACACCAATATGAAGTGGTGAACCATGTTACT 1855  
QY 1913 AGGGCCACAGGCTTTACCGGCCCCACTGGCATCGGACACCATCTACAGATCATATACA 1972  
DB 1856 GAGGCCACCGACTCTACCAAGCGAAGTTGGCGTCCAACTATGTGTATGAGGTGATGCTGA 1915  
QY 1973 GCTGTCGACGAGCTTCAGAAAAAGCGTCCACATTTTCCAGAACCTCTCTCTTCCATTG 2032  
DB 1916 GATGTGCGAGGAGAAACAGGAGGAGGCGCTTCTTTCGAAGATCTGCTGCCACATAG 1975  
QY 2033 AACCACTTCGGGAAAAAGAACAGCAATTTGAAGAGAAATTTAGGAGTG 2078  
DB 1976 ATGAACATAGTGAATGTGAAGAAACTTTTGGGAAGATAAGTGTATGATG 2021

AAS14048

XX ID AAS14048 standard; cDNA; 3593 BP.

XX AC AAS14048;

XX DT 18-DEC-2001 (first entry)

XX DE Human cDNA encoding protein tyrosine kinase, Tec.

XX KW Immunosuppressive; antiinflammatory; osteopathic; antirheumatic;  
XX KW antiarthritic; dermatological; antipsoriatic; antidiabetic; cytostatic;  
XX KW cardiatic; antiarteriosclerotic; vasotropic; cerebroprotective;  
XX KW neuroprotective; nootropic; ophthalmological; immunosuppressive;  
XX KW hepatic; virucide; anti-HIV; immunomodulator; antithyroid; sepsis;  
XX KW septic shock; inflammation; Crohn's disease; rheumatoid arthritis;  
XX KW leukopenia; osteoarthritis; spondyloarthropathy; psoriatic arthritis;  
XX KW severe steroid resistant asthma; pulmonary fibrosis; myocarditis;  
XX KW atherosclerosis; angina; vasculitis; reperfusion injury; infarction;  
XX KW stroke; multiple sclerosis; Alzheimer's disease; Graves ophthalmopathy;  
XX KW psoriasis; transplant rejection; acute alcoholic hepatitis; fibrosis;  
XX KW HIV; human immunodeficiency virus; cachexia; erythema nodosum leprosum;  
XX KW borreliosis; meningococcal septicaemia; cancer; Human; Tec; ss;  
XX KW antibacterial; protein tyrosine kinase.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 57..1913

XX FT /\*tag= a

XX FT /product= "Tec"

XX PN WO200166107-A2.

XX PD 13-SEP-2001.

XX PF 06-MAR-2001; 2001WO-GB00949.

XX PR 06-MAR-2000; 2000GB-0005345.

XX (KENN-) KENNEDY INST RHEUMATOLOGY MATHILDA & TER.

XX PA Foxwell BMJ;

XX PI WP1; 2001-602596/68.

XX DR P-PSDB; AAU09010.

XX Treating condition associated with cytokine production in a mammal,  
XX e.g. sepsis, septic shock, inflammation and Crohn's disease, comprises  
XX administering an inhibitor of a member(s) of the Tec family of protein  
XX tyrosine kinases

XX Claim 23; Fig 13; 78pp; English.

XX The invention relates to the use of inhibitors of the Tec family of  
XX protein tyrosine kinases to treat a condition associated with cytokine  
XX production in a mammal. The Tec family members are e.g. Bruton's  
XX tyrosine kinase (Btk), Tec, Itk, Txk or Bmx. The condition is especially  
XX sepsis, septic shock, inflammation, Crohn's disease or a condition  
XX induced by FRR (Toll related receptor) ligand, LPS (lipopolysaccharide),  
XX zymosan or Gram-negative bacteria. The method may also be used for  
XX treating rheumatoid arthritis. Biological effects of LPS include fever,  
XX circulatory disturbances and vascular hypersensitivity to adrenergic  
XX drugs, leukopenia typically followed by leucocytosis, non-specific  
XX stimulation of B-lymphocytes to undergo blast transformation and  
XX proliferation, lethal toxicity and non-specific tolerance to endotoxin  
XX through repeated exposure to LPS. LPS also causes an increase in  
XX production of cytokines e.g. TNF. Conditions associated with IL-1beta  
XX and TNFalpha osteoarthritis, spondyloarthropathy, psoriatic arthritis,  
XX severe steroid resistant asthma, pulmonary fibrosis, myocarditis,  
XX atherosclerosis, angina with high CRP/IL-6, vasculitis, reperfusion  
XX injury, infarction, stroke, multiple sclerosis, Alzheimer's disease,  
XX Graves ophthalmopathy, psoriasis, transplant rejection, acute alcoholic  
XX hepatitis, fibrosis. Other conditions for which inhibitors of Tec

CC family PTKs may be used include HIV (human immunodeficiency virus  
CC infection, cachexia, erythema nodosum leprosum, borreliosis, or  
CC meningococcal septicaemia, cancer, such as breast cancer, ovarian cancer  
CC or colon cancer. Many other examples of diseases are given in the  
CC specification. The present sequence encodes human Tec.  
XX  
SQ Sequence 3593 BP; 1070 A; 702 C; 865 G; 956 T; 0 other;

Query Match 15.8%; Score 388; DB 22; Length 3593;

Best Local Similarity 59.6%; Pred. No. 9.9e-96;

Matches 674; Conservative 0; Mismatches 450; Indels 6; Gaps 1;

QY 899 AAAACCTGGATGATTATGACGTGGTTTGGTGTGTAACATCTCCAGATACAAATCTGAACAGT 958

DB 775 ACACTTTAGATCAATATGAATGGTATTGCCAGAAATATGAATAGAACGAGCAGACCAAC 834

QY 959 TACTCACACAAAAGGGAAGAACGACGACATTTATGCTTAGAAATTCGAGCCCAAGTGGGAA 1018

DB 835 TCCTCCGACGTGAAGATAAAGAGGTGGTTTATGGTAAGGATTCACAGTCAACAGGCT 894

QY 1019 TGTACACAGTGTCCCTTATTTAGTAAGGCTGTGAATGATAAAAGAACTGTCAACAACT 1078

DB 895 TGTACACAGTCTCCCTTTATACCAAGTTTGGAGGAGAAGTTTCATCGGGTTTTCAGGCAT 954

QY 1079 ACCACGTGCA-----TACAAATGCTGAGAAACAAATATATCTGGCAGAAAACTACTGT 1132

DB 955 ATCATATAAAGGAAACAAACACATCTCCAAAGAAAGTATTACCTAGCTGAAAAACATGCTT 1014

QY 1133 TTGATTCGATTCGAAAGCTTATTCATTTATCATCAACAAATTCAGCAGGCGATGATCACAC 1192

DB 1015 TTGGCTCCATTCCTGAGATTATTGAATATCATTAAGACCAATATGACGAGGACTTGTCAACA 1074

QY 1193 GGCTCCGCCACCTGTGTCAACAAAGGCCCAAGTCCCGGCTCTGTGTCCTCTGGGAA 1252

DB 1075 GGCTTCGGTACCCAGTTAGTGTGAAGGGAAGATGACCCACCCACTCGAGGATTCAGCT 1134

QY 1253 ATGGAATCTGGGAACTGAAAGAGAGAGATTACCTTTTGAAGAGGCTGGGAGTGGCC 1312

DB 1135 ATGAGAAATGGGAGATTAAACCTTTCAGAACTGACCTTTATGAGGGAATTTGGGAAGTGGAC 1194

QY 1313 AGTTTGGAGTGTCCAGCTGGGCAAGTGGGAAGGCGCAGTATGATGTTGCTGTTRAGATGA 1372

DB 1195 TGTTTGGAGTGTGAGGCTTGGCAAAATGGCGAGGCCAGTCAAAAGTTCGCAATCAAAAGTA 1254

QY 1373 TCAAGGAGGCTCCATGTCAAGAGATGAATTTCTTCAGGAGGCCACAGACTATGTAAGAAC 1432

DB 1255 TTCGGAGGTGCAATGTGCGAGGAGGACTTTATAGAAAGAGCTAAAGATGATGATGAAC 1314

QY 1433 TCAGCCATCCCAAGCTGGTTAAATTTCTATGGAGTGTGTTTCAAGAGGAATACCCCATATACA 1492

DB 1315 TGACACACCCGAGTTAGTGCAGCTTTATGGTGTGTCACCCACAGCAAAACCAATATACA 1374

QY 1493 TAGTGACTGAATATATAGCAATGGCTGTGCTGAATATTACTGAGGAGTCCAGGAAAG 1552

DB 1375 TTGTTACTGATTCATGGAAGGGGCTGCCCTTCTGAAATTTCTCCGACAGACAGCAAGGTC 1434

QY 1553 GACTTGAACCTCCACGCTCTAGAAATGCTACGATGCTGTGAAGGCAATGGCTTCT 1612

DB 1435 ATTTTCTAGTAGACAGGTACTGCTGAGCATGTGTGAGGATGTGTGAAGGGATGAGTATC 1494

QY 1613 TGGAGAGTCCCAATTCATACACCGGACTTGGCTGCTCGTAACCTGCTGGTGGACAGAG 1672

DB 1495 TGGAGAAACACGCTTCATCCACAGAGATCTGGCTGCCAGAAATTTGCTAGTAGTGAGG 1554

QY 1673 ATCTCTGTGAAAGTATCTGACTTTGGAAATGACAAGTATGTTTCTTGATGACCAGTATG 1732

DB 1555 CGGAGTTGTAAGAGTATCTGATTTTGAATGGCCAGGTAATGTTCTGATGATGATGATACA 1614

QY 1733 TCAGTTCAGTCGGAACAAAGTTTCCAGTCAAGTGTGTCAGCTCCAGAGGTGTTTCATTA 1792

DB 1615 CAAGTCTTCTGCTGCTGAAGTTTCTGTAAGTGGTGTCCACCTCAAGTGTGTTAATTAACA 1674

QY 1793 TCAATATACACGACGACGAGCTATGGGCAATTTGGGATCTCTGATGTTGGAGGTGTTCA 1852

Db 1675 GCCGCTTCAGCAGCAAAATCAGATGCTGTCTATTTGGTCTTTATATGGGAGATATCA 1734  
 QY 1853 CCCTGGGGAACGACCCCTATGACTTGTATGACAACTCCAGGTGGTCTTGAAGTCTCCC 1912  
 Db 1735 CGSAGGCGAGAAATGCTTTTGAATAATACACCAATATGAGTGGTAACCATGGTTACTC 1794  
 QY 1913 AGGCCACAGCTTTACCGGCCCCACCTGGCATCGGACACCATCTACCAAGATCATGTACA 1972  
 Db 1795 GAGCCACCGACTTACAGCCGGAAGTTGGGCTCCAACTATGTGTATGAGTGTGCTGA 1854  
 QY 1973 GCTGCTGCGACAGCTTCCAGAAAGGCTCCACATTTTCAGCAACTCTCTG 2022  
 Db 1855 GATGTTGCGAGAGAAACAGAGGAGGCTTCTTTGGAAGACTGCTG 1904

## RESULT 15

ABK85703  
 ID ABK85703 standard; DNA; 1368 BP.

XX ABK85703;

XX 20-AUG-2002 (first entry)

XX DNA encoding truncated Tec kinase Itk.

KW Tec kinase; pleckstrin-homology domain; Tec homology domain;  
 KW inflammation; asthma; allergic rhinitis; arthritis; rheumatoid arthritis;  
 KW adult respiratory distress syndrome; osteoarthritis; uveitis;  
 KW inflammatory eye condition; inflammatory bowel condition;  
 KW Crohn's disease; ulcerative colitis; periodontal disease; oesophagitis;  
 KW inflammatory skin condition; psoriasis; eczema; dermatitis; Itk;  
 KW gene; ds; human.

XX Homo sapiens.

XX Key Location/Qualifiers

CDS 7..1362  
 FT /\*tag= a  
 FT /product= "Truncated Itk"  
 FT /note= "Tec kinase"

XX WO200234899-A2.

XX 02-MAY-2002.

XX 17-OCT-2001; 2001WO-EPI1949.

XX 20-OCT-2000; 2000GB-0025804.

XX (GLAX ) GLAXO GROUP LTD.

XX Sims MJ, Hayes DJ;

XX WPI: 2002-471439/50.

XX P-PSDB; AAU98425.

XX New truncated Tec kinase polypeptide useful for identification of a  
 XX compound which modulates the activity of the polypeptide for treating  
 XX inflammatory conditions such as asthma, rheumatoid arthritis, and  
 XX uveitis.

XX Claim 11; Fig 2; 38pp; English.

CC The invention describes a truncated Tec kinase polypeptide (I) having a  
 CC Tec kinase amino acid sequence truncated by a minimum of the amino acids  
 CC constituting the N-terminal pleckstrin-homology (PH) domain and a portion  
 CC of the Tec homology (TH) domain including a proline rich region up to but  
 CC not including the amino acids constituting the kinase domain. A compound  
 CC (II) which modulates Tec activity is useful in treatment of a disorder  
 CC such as inflammation that is responsive to modulation of Tec  
 CC kinase activity. (I) is useful for screening for targets for the  
 CC identification and development of new pharmaceutical agents, including

CC agonists and antagonists of Tec kinases, which may be useful in therapy.  
 CC The nucleic acid (II) encoding (I) is useful for treating inflammatory  
 CC diseases such as asthma, allergic rhinitis, adult respiratory distress  
 CC syndrome, arthritic conditions such as rheumatoid arthritis,  
 CC osteoarthritis, inflammatory eye conditions such as uveitis, inflammatory  
 CC bowel conditions such as Crohn's disease, ulcerative colitis, periodontal  
 CC disease, oesophagitis, and inflammatory skin conditions such as  
 CC psoriasis, eczema and dermatitis. (I) is suitable for screening methods  
 CC as they do not need to be pre-activated by phosphorylation. In vivo, Tec  
 CC kinases need to be phosphorylated by other kinases in order to activate  
 CC the enzyme, while pre-activation by phosphorylation is commonly required  
 CC in assays. The removal of the need to pre-activate (I) offers a  
 CC simplification for the assay. A further advantage of (I) is the provision  
 CC of an assay that is robust as: it is possible to generate large amounts  
 CC of truncated enzyme which are stable over a long time and the assay gives  
 CC a high frequency of comparable results upon repeat testing. This sequence  
 CC encodes a truncated construct of the Tec kinase Itk.

XX Sequence 1368 BP; 370 A; 302 C; 374 G; 322 T; 0 other;

QY Query Match 15.7%; Score 385.4; DB 24; Length 1368;  
 Db Best Local Similarity 59.5%; Pred. No. 3e-95;  
 QY Matches 691; Conservative 0; Mismatches 461; Indels 9; Gaps 2;

QY 839 AAACCTCGATGATTATGACTGGTTGCTGGTAAACATCTCCAGATCACAATCTGAACAGT 958  
 Db 194 ATAATCTGGAAACCTATGAGTGGTACAATAAGAGTATCAGCCGAGACAAAGCTGAAAAAC 253  
 QY 959 TACTCAGACAAAGGAAAGAGGAGCATTTATGGTTAGAAATTCGAGCCAAAGTGGGAA 1018  
 Db 254 TCTTTTGGACACAGGCAAGAGAGGCTTCTCATGGTAAGGATTCAGGAGTGCAGGAA 313  
 QY 1019 TGTACACAGTCTCCTTATTTAGTAAGC---TGTGAATGATAAAAAGGAAGTGTCAAA 1075  
 Db 314 CATACACGCTGCTGTTTTCACCAAGGCTGTTGTAAGTGAGAACAAATCCTGTATAAGC 373  
 QY 1076 ATTACCAGCTGCATCAAAATGCTGA-----GAACAATTTATACCTGGCAGAAAACACT 1129  
 Db 374 ATTATCACATCAAGGAAACAAATGATCAATCTTAAGCGATACCTATGTGGTAAAAAGTATG 433  
 QY 1130 GTTTTGATTCCATCCAAAGCTTATTCATTATCATCAACACAATTCAGCAGCATCATCA 1189  
 Db 434 TGTTGATTCCTATCCCTCTTCTCATCACTATCACAATATGGAGAGGCTGCTGA 493  
 QY 1190 CACGGCTCCGCCCTCTGTGTCAACAAAGGCCAACAAAGTCCCGACTCTGTGTCTCCCTGG 1249  
 Db 494 CTCGACTCCGCTATCCACTTTGTTTGGAGGACAGAAAGCCCACTTACAGCAGGCTGA 553  
 QY 1250 GAAATGGAATCTGGGAAGTGAAGAGAGAGATTTACCTTGTGTGAAGGAGCTGGGAAGTG 1309  
 Db 554 GATACGGGAAATGGGTGATCGACCCCTCAGAGCTCACTTTTGTGCAAGAGATTGGCAGTG 613  
 QY 1310 GCCAGTTTGGAGTGTGCTCCAGCTGGGCAAGTGAAGGGCAGTATGATGTGCTTTAACA 1369  
 Db 614 GCAATTTGGGTGGTGTGCTGCTGGCTACTGGCTCAACAAAGCAAGTGGCTATCAAAA 673  
 QY 1370 TGATCAAGGAGGCTCCATGTGCAAGATGAATTTCTTCAGGAGGCCAGACTATGATCA 1429  
 Db 674 CCAATTCGGAAGGGCTATGTGCAAGAGAGGACTTCATAGAGGAGGCTGAAGTAATGATCA 733  
 QY 1430 AACTCAGCCATCCCAAGCTGTTAAATTTCTATGGAGTGTGTTCAAGGAATACCCCATAT 1489  
 Db 734 AACTCTCTCATCCCAAACTGGTGCAGCTGTATGGGGTGTGCTCGAGCAGGAGGCCCATCT 793  
 QY 1490 ACATAGTGACTGAATATATAGCAATGCTCTCTCTCAATTAACCTAGGAGTCAAGGNA 1549  
 Db 794 GCCTGGGTGTTGATTTTATGGACACCGGCTGCTGTGATATATACGACCCAGCGGG 853  
 QY 1550 AAGGACTTGAACCTTTCCAGCTCTTGAATAATGTGTAGAGTGTCTGTGAAGCATGGCT 1609  
 Db 854 GACTTTTGTGTCAGAGACCCCTGCTGGCATGTGTGTGATGTGTGTGAGGCGATGGCT 913  
 QY 1610 TCTTGGAGAGTCAACAATTCATACACCGGAGCTTGGCTGCTGCTGAATCTGTGTGGACA 1669



```

      | |||
Db 914 ACCTGGAAGGCGATGTGTCATCCACAGAGACTTGGCTGCCAGAAATTTGTTGGTGGAG 973
      | |||
Qy 1670 GAGATCTCTGTGTCAAAGTATCTGACATTTGGAATGACAAGGTATGTTCTTGATGACCACT 1729
      | |||
Db 974 AAAACCAAGTCATCAAGGTGTCTGACATTTGGGATGACAAGGTTCGTTCTGGATGATCAGT 1033
      | |||
Qy 1730 ATGTCAAGTCAGTCGGGAACAAAGTTTCCAGTCAAGTGGTCAGCTCCAGAGAGGTGTTTCATT 1789
      | |||
Db 1034 ACACCAAGTTCCACAGGACCAAAATTCCTGGTGAAGTGGGATCCCCAGAGAGTTTCTCTT 1093
      | |||
Qy 1790 ACTTCAAAATCAGCAGCAAGTCAGACGTATGGGCAATTTGGGATCCTGATGTGGAGGTGT 1849
      | |||
Db 1094 TCAGTCGTATAGCAGCAAGTCCGATGTGTGGTCATTTGGTGTCTGATGTGGGAAGTTT 1153
      | |||
Qy 1850 TCAGCCTGGGNAAGCACCCCTATGACTTGTATGACAACTCCAGGTGGTCTGAAAGTCT 1909
      | |||
Db 1154 TCAGTGAAGGCAAAATCCCTATGAAAACCGAAGCAACTCAGAGGTGGTGAAGACATCA 1213
      | |||
Qy 1910 CCCAGGCCACAGGCTTTACGGGCCCCACCTGGCATCGGACACCATCTACCAAGATCATGT 1969
      | |||
Db 1214 GTACCGGATTCGGTTGTACAAGCCCGGTGGCTCCACACAGCTTACCAAGATTATGA 1273
      | |||
Qy 1970 ACAGCTGCTGGCAGAGCTTCCAGAAAAGCGTCCACATTTTCAGCAACTCCTGTCTTCCA 2029
      | |||
Db 1274 ATCACTGCTGGAAGAGAGAGACCAAGAGATCGGCCAGCCTTCTCCAGAGCTGCTGGGTCAAC 1333
      | |||
Qy 2030 TTGAACCACTTCGGGAAAAG 2050
      | |||
Db 1334 TGGCTGAAATTCAGAAATCAG 1354
```

Search completed: August 20, 2003, 12:44:54  
Job time : 634 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: August 20, 2003, 12:29:37 ; Search time 4600 Seconds  
(without alignments)  
12976.477 Million cell updates/sec

Title: US-10-021-660-41  
Perfect score: 2456  
Sequence: 1 gcaagcagcgacacgctga.....aatgtggaaaaaaacacg 2456

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estov:\*
  - 5: em\_estov:\*
  - 6: em\_estpl:\*
  - 7: em\_estro:\*
  - 8: em\_hic:\*
  - 9: gb\_estl:\*
  - 10: gb\_est2:\*
  - 11: gb\_hic:\*
  - 12: gb\_est3:\*
  - 13: gb\_est4:\*
  - 14: gb\_est5:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: em\_gss\_hum:\*
  - 18: em\_gss\_inv:\*
  - 19: em\_gss\_pln:\*
  - 20: em\_gss\_vrt:\*
  - 21: em\_gss\_fun:\*
  - 22: em\_gss\_mam:\*
  - 23: em\_gss\_mus:\*
  - 24: em\_gss\_pro:\*
  - 25: em\_gss\_rod:\*
  - 26: em\_gss\_png:\*
  - 27: em\_gss\_vrl:\*
  - 28: gb\_gss1:\*
  - 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1641.8	66.8	2866	11 AK040936	AK040936 Mus muscu
2	1636.8	66.6	2864	11 AK036707	AK036707 Mus muscu
3	878.2	35.8	1100	12 BM544910	BM544910 AGENCOURT
4	739.2	30.1	986	12 BM804905	BM804905 AGENCOURT

5	732.8	29.8	830	13	BO878627	BO878627 AGENCOURT
6	729.6	29.7	820	13	BU172565	BU172565 AGENCOURT
7	709	28.9	856	13	BQ229686	BQ229686 AGENCOURT
8	683	27.8	890	13	BQ226991	BQ226991 AGENCOURT
9	635	25.9	887	13	BQ350537	BQ350537 RC3-HT062
10	614.4	25.0	651	13	BQ350538	BQ350538 RC3-HT062
11	592.8	24.1	628	13	BQ350352	BQ350352 RC3-HT062
12	592.2	24.1	688	13	BQ350539	BQ350539 RC3-HT062
13	569	24.0	1014	11	AK080038	AK080038 Mus muscu
14	572.4	23.3	1029	12	BM542399	BM542399 AGENCOURT
15	553	22.5	590	10	BE180268	BE180268 RC3-HT062
16	530.4	21.6	805	14	CB998198	CB998198 AGENCOURT
17	504.4	20.5	665	10	BB629678	BB629678 BB629678
18	458.4	18.7	554	10	BE896015	BE896015 601438807
19	443.6	18.1	962	13	BQ226370	BQ226370 AGENCOURT
20	438	17.8	465	13	BX102696	BX102696 BX102696
21	428	17.4	670	10	BB637076	BB637076 BB637076
22	424.6	17.3	680	10	BB218897	BB218897 BB218897
23	424.4	17.3	458	12	BG979101	BG979101 CM4-CN009
24	420.2	17.1	667	10	BB636992	BB636992 BB636992
25	419	17.1	2564	11	BC018230	BC018230 Mus muscu
26	406.6	16.6	449	9	AI335555	AI335555 ql30e06.x
27	398.4	16.2	4208	11	AK083380	AK083380 Mus muscu
28	392	16.0	2188	11	AK088209	AK088209 Mus muscu
29	370.4	15.1	896	10	BE894841	BE894841 601434022
30	360	14.7	380	9	AI500626	AI500626 tn98903.x
31	359.6	14.6	4155	11	AK037554	AK037554 Mus muscu
32	356	14.5	1033	13	BQ429846	BQ429846 AGENCOURT
33	346	14.1	1355	13	BQ714415	BQ714415 AGENCOURT
34	344.8	14.0	961	13	BQ461627	BQ461627 603370561
35	344	14.0	533	10	BF898833	BF898833 IL0-WT028
36	335.8	13.7	459	9	AA187327	AA187327 zp60a08.s
37	323.4	13.2	533	10	BE749703	BE749703 200748 MA
38	315.4	12.8	1039	12	BM478046	BM478046 AGENCOURT
39	308	12.5	328	9	AI382692	AI382692 tc61b08.x
40	304.4	12.4	478	10	BE623796	BE623796 ut22h09.y
41	289	11.8	312	14	240071	240071 HSC1Q6122.n
42	283	11.5	288	9	AW512350	AW512350 xx73e10.x
43	282.2	11.5	962	10	BF795127	BF795127 602256462
44	277.2	11.3	723	9	AJ395612	AJ395612 AJ395612
45	271.8	11.1	758	14	CB310787	CB310787 AGENCOURT

## ALIGNMENTS

RESULT 1  
AK040936  
LOCUS  
DEFINITION  
AK040936 2986 bp mRNA linear HTC 05-DEC-2002  
Mus musculus adult male aorta and vein cDNA, RIKEN full-length  
enriched library, clone:A530048A03 product:BMX non-receptor  
tyrosine kinase, full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AK040936  
AK040936.1 GI:26334070  
HTC: CAP trapper.  
Mus musculus (house mouse)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
1  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
PUBMED 11042159



QY	547	AGATACCTCGGCGAGTTCCCTGTTCTCAAAATGGATGACCATCTTCAAGTACCACCTCTA	606
Db	670	AAAGTCCAGGGCAGTTCCCGTCTTCAAAATGGATGACCATCTTCAAGTACCACCTCTA	729
QY	607	GCCCAATATGACAAAGATCAAGAAATCAAGAAATCAAGTCCAGCCACCATCTTCAAGTACC	666
Db	730	CCC-----	732
QY	567	AGTCTAGCGCAATATGACAGCAACTCAAGAAATCTATGGCTCCAGCCCAACATCTCAAC	726
Db	733	-----CAATATGACAGCTATTTCAAGAAAGTGTGGTTCCAGCCCAACAGCAAC	783
QY	727	ATGCAGTATATCCAAAGGAGACTTCCCTGACTGGTGGCAAGTAAAGAACTGAAAGT	786
Db	784	ATACGCTATATCCAAAGGAGACTTCCCTGACTGGTGGCAAGTAAAGAACTGAAAGT	840
QY	787	AGCAGCAGCTGAAGATGTTGCAAGCAGTAAACCAAAAGAAAGAAAT---GTGAATCAC	843
Db	841	-----ACGAGGAGAGACATTTGCATGCAGTAAACCACTGGAAAGAAATATGCCCTCTCAC	894
QY	844	ACCACCTCAAGATTTCAATGGAAATTCCTCAGTCAAGTTCATCTGAAAGAGAGAAAC	903
Db	895	AGCACCTCAAGATGTCATGGGATTCCTCAGTCAAGTTCATCAGAAGAGAGAAAT	954
QY	904	CTGATGATATGACTGGTTTCTGGTGAACATCTCCAGATCACAACTGACAGATTACTC	963
Db	955	CTGATGATATGACTGGTTTCTGGGAAATATCTCCAGGTACAACTGAGCAGTTACTG	1014
QY	964	AGACAAAGGAGAAAGAGAGCATTTATGTTAGAAATTCGAGCAAGTGGGAATGTAC	1023
Db	1015	AGACAAAGGAGAAAGAGAGCATTTATGTTGGAATTCGAGCAAGTGGGAATGTAC	1074
QY	1024	ACAGTGCTCTTATTTAGTAAAGCTGTGAATGATAAAAGAACTGTCAAACTATTACCAC	1083
Db	1075	ACTGTGCTCTTATTTAGTAAAGCTGTGAATGATAAAAGAACTGTCAAACTATTACCAC	1134
QY	1084	GTGCATACAAATGTGAGAAACAATATACCTGGCAGAAACTACTGTTTGTGATTCCAAT	1143
Db	1135	GTGCATACAAATGTGAGAAACAATATACCTGGCTGAAACTACTGTTTGTGATTCCAAT	1194
QY	1144	CCAAAGCTTATTCATTAACACAAATTCAGCAGGCATGATCACACGGCTCCGCCAC	1203
Db	1195	CCAAAGCTTATTCATTAACACAAATTCAGCAGGCATGATCACACGGCTCCGCCAC	1254
QY	1204	CCTGTGCAACAAAGGCCAACAAGTCCCGACTCTGTCCCTGGGAAATGGAATCTGG	1263
Db	1255	CCAGTGTCAACCAAGGCCAACAAGTCCCGGTCTGTGGCTCTGGGAAGTGGAAATGG	1314
QY	1264	GAACTGAAAGAGAGAGATTTACCTGTTGAAGAGCTGGGAAGTGGCCAGTTGGAGTG	1323
Db	1315	GAACTGAAAGAGAGAGATTTACCTGTTGAAGAGCTGGGAAGTGGCCAGTTGGAGTG	1374
QY	1324	GTCCAGCTGGGCAAGTGAAGGGCAGTATGATGTTGCTGTTAAGATGATCAAGAGGGC	1383
Db	1375	GTCCAGCTGGGCAAGTGAAGGGCAGTATGATGTTGCTGTTAAGATGATCAAGAGGGT	1434
QY	1384	TCCATGTCAAGAGATGAATTTCTTTCAGGAGGCCAGACTATGATGAAACTCAGCCATCCC	1443
Db	1435	GCCATGTCAAGAGATGAATTTCTTTCAGGAGGCCAGACTATGATGAAACTCAGCCATCCC	1494
QY	1444	AAGCTGGTTAAATTTCTATGGAGTGTGTTCAAGGAAATACCCCATATACATAGTACTGAA	1503
Db	1495	AAGCTGGTTAAATTTCTATGGAGTGTGTTCAAGGAAATACCCCATATATAGTACTGAG	1554
QY	1504	TATATAAGCAATGCTGCTTCTGAAATTAACCTGAGGAGCTACGGAAGAGACTTGAACCT	1563
Db	1555	TATATAAGCAATGCTGCTTCTGAAATTAACCTGAGGAGCTACGGAAGAGACTTGAAGT	1614
QY	1564	TCCAGCTCTTAGAAATCTAGCATGCTGTGAAGGCATGGCCCTTCTTGGAGAGTCCAC	1623
Db	1615	TGCCAGCTCTTAGAAATCTAGCATGCTGTGAAGGCATGGCCCTTCTTGGAGAGCTCAT	1674
QY	1624	CAATTATACATACCGGACCTTGGCTGCTAACTGCTGTCGTCGACAGAGATCTCTGTGTG	1683
Db	1675	CAGTTATACATACCGGACCTTGGCTGCTGCGAACTGTTGGTGACAGATGATCTCTGTGTG	1734
QY	1684	AAAGTATCTGACCTTTTGGAAATGACAAAGTATGTTCTTTGATGACCAAGTATGTCAGTCT	1743
Db	1735	AAAGTCTCAGACTTTTGGAAATGACAGAGATATGCTTTGATGATCAGTATGTCAGTCTCAGTA	1794
QY	1744	GGAACAAAGTTTCCAGTCAAGTGGTCCAGCTCCAGAGGTTTTCATTACTTCAAAATACAGC	1803
Db	1795	GGAACAAAGTTTCCAGTCAAGTGGTCCAGCTCCAGAGGTTTTCATTACTTCAAAATACAGC	1854
QY	1804	AGCAAGTCAAGCTATGGGCAATTTGGATCCCTGATGTCGGAGCTGTCAGCTCGGGAAG	1863
Db	1855	AGCAAGTCCGATGATGGGCAATTTGGATCCCTGATGTCGGAGCTGTTAGTCTAGGGAAG	1914
QY	1864	CAGCCCTATGACTTTGATGACAACTCCAGGTGGTTCTGAAGTCTCCCAAGGCCACAGG	1923
Db	1915	CAGCCCTATGACTTTATGATGAATCTCCAGGTGGTTCTGAAGTCTCCCAAGGCCACAGG	1974
QY	1924	CTTTACCGGCCCACTGGCATCGGACACCATCTACAGATCATGTACAGCTGCTGGCAC	1983
Db	1975	CTCTACCGGCCCACTGGCATCGGACACCATCTACAGATCATGTACAGCTGCTGGCAT	2034
QY	1984	GAGCTTCCAGAAAGGGTCCACATTTTCAGCAACTCTCTGCTCCCATAGAACCTTCGG	2043
Db	2035	GAGCTTCCAGAAAGGGTCCACATTTTCAGCAACTCTCTGCTCCCATAGAACCTTCGG	2094
QY	2044	GAAAAAGCAACGATTTGAAGAAATTAGGAGTCTGTAAGAAATGAATATAGATGCTG	2103
Db	2095	GAAACAGATAAACCATGAAGAGAAATTTTCAGAGTGGTATGAGAAACAATGTCATATTA	2154
QY	2104	GCCAGCAATTTTCATTCATTTTAAAGAAAGTGAAGAGCAATAAGTAAATTTTAGTACTGTT	2163
Db	2155	ACTAACACCTCTCATTTGTTGAAGGAAGTGAAGAGACTCAATGCAATGTAGTCTGCT	2214
QY	2164	TAATAGTGTCTCTGATTTGCTATTTATTTA-----CAAAATGAACAGGACGGA	2212
Db	2215	TAACATGCTCTATGCTGCTTTATCTATCTACCAAAAAAGCAAGGTAGGA	2274
QY	2213	AACAAAGATTTCCCTTGAATTTAGATCAAAATTTAGTAAATTTGTTTATGCTGCTCTGA	2272
Db	2275	AACAGATAGTCTTTGAAATTTAGTCAAAATTTGTTGTTGCTGCTCTTTAA	2333
QY	2273	TATAACATTTCCAGCTATAGCAGAACACATTTTCAGACTGCAATATAGAGACTGTG	2332
Db	2334	GST-ATACTTCCAAGACTATGTTAGAAATTTAGTCAAAATTTAGTCAAAATTTGCTAC	2392
QY	2333	TCATGTGTAAGACTGAGCAGAACGAAATTTACTTTTATGATATTCATTCTTTCT--	2390
Db	2393	T-GTGTATAAAGGTAGAGCAGAAATTTGAAAGGTGATTTGTTGGAACTCCCTTTTAA	2451
QY	2391	---TTATATTGCTATGCTCACAAATTTAAATTAATTAATTAATTAATTAATTAATTAAT	2441
Db	2452	AAAAAATGGCAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2504
RESULT 2			
AK036707			
LOCUS	AK036707		
DEFINITION	Mus musculus adult male bone cDNA, RIKEN full-length enriched library, clone:9830165F06 product:BMX non-receptor tyrosine kinase, full insert sequence.		
ACCESSION	AK036707		
VERSION	AK036707.1 GI:26331623		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1 Carninci, P. and Hayashizaki, Y.		
AUTHORS	High-efficiency full-length cDNA cloning		
TITLE			

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Meth. Enzymol. 303, 19-44 (1999)

99279253  
10349636

2

TITLE

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

20493374

11042159

3

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

REFERENCE

AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staehli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, F., Furuno, M., Aono, H., Baldarelli, R., Barsh, C., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., De Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Momberti, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 2864)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saiton, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tanaka, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001)

Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp).

COMMENT

URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.go.jp/>  
URL: <http://genome.gsc.riken.go.jp/>

FEATURES

source

1. .2864  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:9830165F06"  
/db\_xref="taxon:10090"  
/clone="9830165F06"  
/sex="male"  
/tissue\_type="bone"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
15. .1970  
/note="unnamed protein product; BMX non-receptor tyrosine kinase (MGDI|MG1:1101778, GB|NM\_009759, evidence: BLASTN, 99%, match=2607)  
putative"  
/codon\_start=1  
/protein\_id="BAC29542.1"  
/db\_xref="GI:26331624"  
translation="MESKLELELLKKKSKKSPNNYKERFLVLTSLSYEYD KMRGSRGSTEIKRCVKNVLEOTPVQYPOIVYKDGILLVYVSNESRCOW LKALQKIRGNPHLLIYHSGFFVDGKFLCCQSCKAPGCTLAEAYADLHAI SDEK HRAPTFFERLLKIPRAVPLKMDASSGAILPQDSYKSKCGSGTNSINIRYIPRED C PDWQVRLKSEEDIACSNOLERNIAHSISMVSSSESEENHAYDWFAGN ISRSQQLLRKQKGAIVRNSSQMGMTVSLFSAVNDKGTGTVRYHYVHTNAENK LYLAENYCFDSIPKLIHYHONSAMGTRLRHPVSTKANKVPVSVAGLSGLWELKREE ITLKGNGQFVVGOLQWGMKIDVAKMIKEGAMSEDEFQEAOTWMLSHPLKLVK FYGVCSEKPIYIVTEYITNGCLLYLKHGKLESCOLLEMCYDVCRCMAFLSHQF IHRDLAARNLVDSLSVYKSDVDFNRYLDDQYVSSVGTGTFPKVNSAPEVPHYKYS WHSDVAFGLMVEVFLSDGKPYDLYDNSEVVVVKVYSGQHRILYRPLASDTIYQIMISC WHELPERPTFOQLLSAIEPLREQDKP"

CDS

polyA\_signal  
polyA\_site  
BASE COUNT 877 a 564 c 674 g 749 t  
ORIGIN

Query Match 66.6%; Score 1636.8; DB 11; Length 2864;  
Best Local Similarity 82.5%; Pred. No. 0;  
Matches 2012; Conservative 0; Mismatches 332; Indels 96; Gaps 8;  
Qy 20 AGACGGATGATAATGATGATACAAAATCTATCTAGAGAACTCTCTCTCAAAAGATCAC 79  
Db 1 AAATGGACGCAATATGGAGAGTAATCCATCTCTGGAAGAACTGCTCTCAAAAGATCAC 60  
Qy 80 AGCAAAAAGAGAAAATGTACCAAAATATTTACAAAGACGGCTTTTGTGACCAAAA 139  
Db 61 AGCAAAAAGAGAAAATGTACCAAAATATTTACAAAGACGGCTTTTGTGACCAAAA 120  
Qy 140 CAAACCTTCTCTACTATGATATGACAAAATGAAAGGGGAGAGAGATGATGATG 199  
Db 121 CAAGCCCTTCTCTACTATGATATGACAAAATGAAAGGGGAGAGAGATGATGATG 180  
Qy 200 AAATTAAGAAAATCAGATGTGTGAGAGAAAGTAAATCTCGAGGAGCAGACGGCTGTAGAGA 259  
Db 181 AAATTAAGAAAATCAGATGTGTGAGAGAAAGTAAATCTCGAGGAGCAGACGGCTGTAGAGA 240  
Qy 260 GACAGTACCCATTTTCAGATGTGTATTAAGATGGGCTTCTCTATGCTATGCTATGCTATGCTATG 319  
Db 241 GACAGTACCCATTTTCAGATGTGTATTAAGATGGGCTTCTCTATGCTATGCTATGCTATGCTATG 300







found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12716 row: m column: 21  
High quality sequence stop: 644.  
Location/Qualifiers  
1. 986

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5725628"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_125"  
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;  
Site\_1: EcoRV (destroyed); Site\_2: NotI: RNA source pool  
of three ovaries, from females ranging in age from 38 to  
49 yo. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 2.1 kb, insert size range 1-3.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 036." 306 a 193 c 246 g 240 t 1 others

BASE COUNT 306 a 193 c 246 g 240 t 1 others  
ORIGIN

Query Match 30.1%; Score 739.2; DB 12; Length 986;  
Best Local Similarity 98.5%; Pred. No. 3.8e-138;  
Matches 778; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

QY 782 AAGTAGCAGCAGCAGTGAAGATGTTGCAAGCAGTAAACCAAGAAAGAAATGTGAATC 841  
DB 7 AGAGTAGCAGCAGCAGTGAAGATGTTGCAAGCAGTAAACCAAGAAAGAAATGTGAATC 66  
QY 842 ACACCACCTCAAGATTTTCATGGGAATTCCTGAGTCAAGTTCAATCTGAAGAAGAGGAAA 901  
DB 67 ACACCACCTCAAGATTTTCATGGGAATTCCTGAGTCAAGTTCAATCTGAAGAAGAGGAAA 126  
QY 902 ACCTGGATGATTATGACTGTTTCTGTTACATCTCCAGATCAATCTGAACAGTTAC 961  
DB 127 ACCTGGATGATTATGACTGTTTCTGTTACATCTCCAGATCAATCTGAACAGTTAC 186  
QY 962 TCACAGAAAAGGAAAAGAGGAGCATTATGTTAGAAATTCGAGCAAGTGGGAATGT 1021  
DB 187 TCACAGAAAAGGAAAAGAGGAGCATTATGTTAGAAATTCGAGCAAGTGGGAATGT 246  
QY 1022 ACACAGTGTCTTATTTAGTAAAGCTGTGAATGATATAAAAGGAACCTGTCAAACTTACC 1081  
DB 247 ACACAGTGTCTTATTTAGTAAAGCTGTGAATGATATAAAAGGAACCTGTCAAACTTACC 306  
QY 1082 ACCTGCATACAAATGCTGAGAACAAATTTATACCTGGCAGAAACTACTGTTTTCATTTCCA 1141  
DB 307 ACCTGCATACAAATGCTGAGAACAAATTTATACCTGGCAGAAACTACTGTTTTCATTTCCA 366  
QY 1142 TTCAAAGCTTATTCATTATCATCAACAAATTCAGCAGGCATGATCACAGGCTCCGCC 1201  
DB 367 TTCAAAGCTTATTCATTATCATCAACAAATTCAGCAGGCATGATCACAGGCTCCGCC 426  
QY 1202 ACCTGTGTCAACAAAGCCCAACAGTCCCGACTCTGTCTCCCTGGGAATGGAATCT 1261  
DB 427 ACCTGTGTCAACAAAGCCCAACAGTCCCGACTCTGTCTCCCTGGGAATGGAATCT 486  
QY 1262 GGGAACTGAAAAGAGAGATTAACCTGTTGAAGAGCTGGGAAGTGGCCAGTTTGGAG 1321  
DB 487 GGGAACTGAAAAGAGAGATTAACCTGTTGAAGAGCTGGGAAGTGGCCAGTTTGGAG 546  
QY 1322 TGCTCCAGCTGGGCAAGTGAAGGGGCGAGTATGATGTTGCTGTTAAGATGATCAAGGAGG 1381  
DB 547 TGCTCCAGCTGGGCAAGTGAAGGGGCGAGTATGATGTTGCTGTTAAGATGATCAAGGAGG 606  
QY 1382 GCTCCATGTCAGAGATGAATTTTTCAGAGGCCCACTATGATGAATCAAGCCATC 1441  
DB 607 GCTCCATGTCAGAGATGAATTTTTCAGAGGCCCACTATGATGAATCAAGCCATC 666  
QY 1442 CCAGCTGCTTAATTTCTATGGAGTGTGTTCAAGGAATACCCCATATACATAGTAGTACTG 1501

DB 667 CCAAGCTGTTAAATCTATGAGAGTGTGTTCAAGGAATACCCCATATACATAGTAGTACTG 726  
QY 1502 AATATATAAGCAAT-GGCTGCTTGTCTG-AAATACCTGAGGAGTCAAGG-AAAAGGACTT 1557  
DB 727 AATATATAAGCAATGGCTGCTGCTGTAATACCTGAGAGTCAAGGGAAGGAGGACTT 786  
QY 1558 GAACCTTCCC 1567  
DB 787 GAACCTTCCC 796

RESULT 5  
BQ878627

LOCUS BQ878627 830 bp mRNA linear EST 16-AUG-2002  
DEFINITION AGENCOURT\_8124174 Lupski\_dorsal\_root\_ganglion Homo sapiens cDNA  
clone IMAGE:6178136 5', mRNA sequence.

ACCESSION BQ878627  
VERSION BQ878627.1 GI:22270635  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 830)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13557 row: d column: 09  
High quality sequence stop: 707.  
Location/Qualifiers  
1. 830

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6178136"  
/sex="male"  
/tissue\_type="dorsal root ganglia"  
/dev\_stage="adult, 36 yr"  
/lab\_host="DH10B"  
/clone\_lib="Lupski\_dorsal\_root\_ganglion"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
NotI; Site\_2: SalI; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCGACCCACGCTCCG-3' and  
5'-GACTAGTCTAGATCGCAGGCGCCCT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies."

BASE COUNT 284 a 189 c 181 g 176 t  
ORIGIN

Query Match 29.8%; Score 732.8; DB 13; Length 830;  
Best Local Similarity 99.6%; Pred. No. 7.2e-137;  
Matches 745; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 GCAAGCAGGGAACAGCTGAGCAGGATGATATATGATACAAATCTATTCTAGAGAA 60  
DB 75 GCAAGCAGGGAACAGCTGAGCAGGATGATATATGATACAAATCTATTCTAGAGAA 134  
QY 61 CTTCTTCTCAAAAGATCACACCAAGAAAGAAATGTCCACCAATATATACAAAGACGG 120

```

135 CTTCTCTCAAAAGATCACAGCAAAAGAAAGAAATGTCACCAATATATTACAAAGACGG 194
121 CTTTTTGTGTTTGACCAAAACAAACCTTTCTACTATGAATATGACAAAATGAAAAGGGG 180
195 CTTTTTGTGTTTGACCAAAACAAACCTTTCTACTATGAATATGACAAAATGAAAAGGGG 254
181 AGCAGAAAAGGATCCATTTGAATTAAGAAATAGAAATCAGATGCTGCGAAGTAAATCTCGAG 240
255 AGCAGAAAAGGATCCATTTGAATTAAGAAATCAGATGCTGCGAAGTAAATCTCGAG 314
241 GAGCAGAGCGCTGTAGAGAGACAGTACCCATTTCAGATTGCTATATAAGATGGCTTCTC 300
315 GAGCAGAGCGCTGTAGAGAGACAGTACCCATTTCAGATTGCTATATAAGATGGCTTCTC 374
301 TATGCTATGCATCAATGAAGAGAGCGGAGTCAAGTCAAGTGGTGAAGCATTACAAAAGAG 360
375 TATGCTATGCATCAATGAAGAGAGCGGAGTCAAGTCAAGTGGTGAAGCATTACAAAAGAG 434
361 ATAAGGGGTAAACCCACCTGCTGGTCAAGTACCATAGTGGTTCCTGCGAGCGGAAG 420
435 ATAAGGGGTAAACCCACCTGCTGGTCAAGTACCATAGTGGTTCCTGCGAGCGGAAG 494
421 TTCTCTGTGTTGCCAGCAGAGCTGTAAAGCAGAGCGCCAGGATGTACCCCTCTGGGAGCATAT 480
495 TTCTCTGTGTTGCCAGCAGAGCTGTAAAGCAGAGCGCCAGGATGTACCCCTCTGGGAGCATAT 554
481 GCTAATCTGCATCTACTGAGTCAATGAAGAGAAACACAGAGTTCCTCCACTTCCAGACAGA 540
555 GCTAATCTGCATCTACTGAGTCAATGAAGAGAAACACAGAGTTCCTCCACTTCCAGACAGA 614
541 GTGCTGAAGATACCTCGGCGAGTTCCTGTTCTCAAAATGATGCACCATCTCAAGTACC 600
615 GTGCTGAAGATACCTCGGCGAGTTCCTGTTCTCAAAATGATGCACCATCTCAAGTACC 674
601 ACTCTAGCCCAATATGACAAAGTCAAGTCAAGAAAGTATGGCTCCAGCCAGCATCTTCA 660
675 ACTCTAGCCCAATATGACAAAGTCAAGTCAAGAAAGTATGGCTCCAGCCAGCATCTTCA 734
661 AGTACCGTCTAGGCGAATATGACAGCACTCAAGAAATCTATGGCTCCAG-CCAAA 719
735 AGTACCGTCTAGGCGAATATGACAGCACTCAAGAAATCTATGGCTCCAGCCAGCA 794
720 CTTCAACATCAGTATATTTCCAGGGAA 747
795 CTTCAACATCAGTATATTTCAAGGGGA 822

RESULT 6
LOCUS BUI72565
DEFINITION BUI72565 820 bp mRNA linear EST 04-SEP-2002
AGENCOURT_7960517 NIH_MGC_72 Homo sapiens cdna clone IMAGE:6166463
5', mRNA sequence.
BUI72565
VERSION BUI72565.1 GI:22686549
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 820)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTP
cdna Library Preparation: Life Technologies, Inc.
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13526 row: m column: 24

```

```

FEATURES
source
High quality sequence stop: 614.
Location/Qualifiers
1..820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6166463"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 236 a 166 c 211 g 205 t 2 others
ORIGIN
Query Match 29.7%; Score 729.6; DB 13; Length 820;
Best Local Similarity 99.1%; Pred. No. 3.2e-136;
Matches 765; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
QY 1059 AAAAGGAACCTGTCAACATTTACACGTCGATACAAATGCTGAGAACAAATATATACC-TGG 1117
DB 1 AAAAGGAACCTGTCAACATTTACACGTCGATACAAATGCTGAGAACAAATATATACC-TGG 60
QY 1118 CAGAAACTACTGTTTGTATTCATTCACAGCTTATTCATTCATCAACACAAATTCAG 1177
DB 61 CAGAAACTACTGTTTGTATTCATTCACAGCTTATTCATTCATCAACACAAATTCAG 120
QY 1178 CAGCATGATCACAGCGCTCGCCACCCCTGTGTCAACAAAGGCCAACAGGTCCCGGACT 1237
DB 121 CAGCATGATCACAGCGCTCGCCACCCCTGTGTCAACAAAGGCCAACAGGTCCCGGACT 180
QY 1238 CTGTGTCCTCGGAAATGGAATCTGGAATCTGGAACCTGGAAGAGAGAGATTAACCTTGTGAAGG 1297
DB 181 CTGTGTCCTCGGAAATGGAATCTGGAACCTGGAAGAGAGAGATTAACCTTGTGAAGG 240
QY 1298 AGCTGGGAAGTGGCCAGCTTTGAGCTGGTCCAGCTGGGCAAGTGGGAAGGGGAGTATGATG 1357
DB 241 AGCTGGGAAGTGGCCAGCTTTGAGCTGGTCCAGCTGGGCAAGTGGGAAGGGGAGTATGATG 300
QY 1358 TTGCTGTTAGATGATCAAGAGGGCTCCATGTCCAGAGATGAATTTCTTTTCAGGAGGCC 1417
DB 301 TTGCTGTTAGATGATCAAGAGGGCTCCATGTCCAGAGATGAATTTCTTTTCAGGAGGCC 360
QY 1418 AGACTATGATAAACTCAGCCATCCCAAGCTGGTAAATTTCTATGAGAGTGTGTTCAAAGG 1477
DB 361 AGACTATGATAAACTCAGCCATCCCAAGCTGGTAAATTTCTATGAGAGTGTGTTCAAAGG 420
QY 1478 AATACCCCATATACATAGTACTGAATATATAAGCAATGGTGTGCTGTAATTAACCTGA 1537
DB 421 AATACCCCATATACATAGTACTGAATATATAAGCAATGGTGTGCTGTAATTAACCTGA 480
QY 1538 GGAGTCACGGAAAAGGACTTGAACCTTCCAGCTTTAGAAAATGTGTACAGATGTCTGTG 1597
DB 481 GGAGTCACGGAAAAGGACTTGAACCTTCCAGCTTTAGAAAATGTGTACAGATGTCTGTG 540
QY 1598 AAGCATGGCTTCTTTGAGAGTCAACAAATTCATACCCGGGACTTGGCTGCTGAACCT 1657
DB 541 AAGCATGGCTTCTTTGAGAGTCAACAAATTCATACCCGGGACTTGGCTGCTGAACCT 600
QY 1658 GCTTGGTGGACAGAGATCTGTGTGAAGTATCTGACTTTGGAAATGACAAGGTATGTC 1717
DB 601 GCTTGGTGGACAGAGATCTGTGTGAAGTATCTGACTTTGGAAATGACAAGGTATGTC 660
QY 1718 TTGATGACCACTATGCTAGTTCAGTTCAGTCGGAACAAAGTTCCTCAAGTCAAGTGGTCAAG 1776
DB 661 TTGATGACCACTATGCTAGTTCAGTTCAGTCGGAACAAAGTTCCTCAAGTCAAGTGGTCAAG 720
QY 1777 GAGGTG-TTTTCAATTAATACAGCAGCAAGTTCAGAGTATGGGCAATTT 1827
DB 721 GAGGTG-TTTTCAATTAATACAGCAGCAAGTTCAGAGTATGGGCAATTT 772

```





```

QY 1710 GTATGTTCTTGATGACGACGATGTCAGTTCAGTCGGACAAAGTTTCCAGTCAAGTGTC 1769
Db 147 GTATGTTCTTGATGACGACGATGTCAGTTCAGTCGGACAAAGTTTCCAGTCAAGTGTC 88
QY 1770 AGTCCAGAGTGTTTCTTACTTCAATACAGCAGCAGCAGTATGGGATTGG 1829
Db 87 AGTCCAGAGTGTTTCTTACTTCAATACA-CAGCAAGTCAGACGT-ATTGGCATTGG 30
QY 1830 GATCCTGATGTGGGA 1844
Db 29 GATCCTAATGTTAGA 15

RESULT 10
BQ350538/c
LOCUS RC3-H70622-130400-022-f06 HT0622 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ350538
VERSION BQ350538.1 GI:21014594
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 651)
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3st2-RC3-HT0622-
130400-022-f06&tl=2000-04-13&tl=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 613.
FEATURES
Location/Qualifiers
source
1..651
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0622"
/notes="Organ: head,neck; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORFESTES PCR (U.S. Letters patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 167 a 176 c 130 g 178 t
ORIGIN

Query Match 25.0%; Score 614.4; DB 13; Length 651;
Best Local Similarity 98.8%; Pred. No. 4.6e-113;
Matches 640; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 1194 GTCGCCGCCACCTGTGTCAACAAAGCCACACAGGTCCCGACCTCTGTGTCCTGGGAAA 1253

```

```

Db 651 GCGCGCCACCCCTGTGTCAACAAAGCCCTACAAGGTCCCGACTCTGTCTCCCTGGGAAA 592
QY 1254 TGGAACTCTGGGAAGTGAAGAGAGAGATTTACCTTTGTTGAAGGAGCTGGGAAGTGGCCA 1313
Db 591 TGGAACTCTGGGAAGTGAAGAGAGAGATTTACCTTTGTTGAAGGAGCTGGGAAGTGGCCA 532
QY 1314 GTTTGGAGTGGTCCAGCTGGGCAAGTGGGAAGGGGAGTATGATGTTGCTGTTAAGATGAT 1373
Db 531 GTTTGGAGTGGTCCAGCTGGCAAGTGGGAAGGGGAGTATGATGTTGCTGTTAAGATGAT 472
QY 1374 CAAGGAGGCTCCATGTCAGAGATGAATTTCTTTCAGGAGGCCAGACATATGATGAACCT 1433
Db 471 CAAGGAGGCTCCATGTCAGAGATGAATTTCTTTCAGGAGGCCAGACATATGATGAACCT 412
QY 1434 CAGCCATCCCAAGCTGGTTAAATTCATGAGAGTGTGTTCAAGAGATATCCCATATACAT 1493
Db 411 CAGCCATCCCAAGCTGGTTAAATTCATGAGAGTGTGTTCAAGAGATATCCCATATACAT 352
QY 1494 AGTGACTGAATATATAAGCAATGGCTGTGCTGAATACCTGAGGAGTCACGGAAAGG 1553
Db 351 AGTGACTGAATATATAAGCAATGGCTGTGCTGAATACCTGAGGAGTCACGGAAAGG 292
QY 1554 ACTTGAACCTTCCAGCTCTTAGAAATGCTGTACGATGTCTGTGAAGGATGCGCTTCTT 1613
Db 291 ACTTGAACCTTCCAGCTCTTAGAAATGCTGTACGATGTCTGTGAAGGATGCGCTTCTT 232
QY 1614 GGAGAGTCACCAATTCATACACCGGAGTGTGCTGCTCGTAACTGCTGGTGGACAGAGA 1673
Db 231 GGAGAGTCACCAATTCATACACCGGAGTGTGCTGCTCGTAACTGCTGGTGGACAGAGA 172
QY 1674 TCTCTGTGTGAAGTATCTGACTTTGGAATGACAGAGTATGTTCTTGATGACCAAGTATGT 1733
Db 171 TCTCTGTGTGAAGTATCTGACTTTGGAATGACAGAGTATGTTCTTGATGACCAAGTATGT 112
QY 1734 CAGTTCAGTCGGACAAAGTTTCCAGTCAAGTGGTCCAGAGTGTCTTCACTTACTT 1793
Db 111 CAGTTCAGTCGGACAAAGTTTCCAGTCAAGTGGTCCAGAGTGTCTTCACTTACTT 52
QY 1794 CAATATACAGCAGCAGTATGAGGATGCGATTTGGGATCCTGATGTG 1841
Db 51 CAATATACAGC-GCAAGTCAGACGTAT-GGCATTTGGGATCCTTAACATG 6

```

```

RESULT 11
BQ350352 628 bp mRNA linear EST 20-MAY-2002
LOCUS RC3-HT0622-010400-021-b05 HT0622 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ350352
VERSION BQ350352.1 GI:21014408
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 628)-
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922

```



Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?td=RC3st2-RC3-HF0622-010400-021-b05st3-2000-04-01st4-1>)  
Seq primer: puc 18 forward  
High quality sequence start: 11  
High quality sequence stop: 537.

RESULT 12  
BQ350539/c

BQ330339/C	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
------------	-------	------------	-----------	---------	----------	--------	----------

ORGANISM  
.  
REFERENCE  
AUTHORS

AUTHORS	TITLE	JOURNAL MEDLINE PUBMED COMMENT
1	2	3
4	5	6
7	8	9
10	11	12
13	14	15
16	17	18
19	20	21
22	23	24
25	26	27
28	29	30
31	32	33
34	35	36
37	38	39
40	41	42
43	44	45
46	47	48
49	50	51
52	53	54
55	56	57
58	59	60
61	62	63
64	65	66
67	68	69
70	71	72
73	74	75
76	77	78
79	80	81
82	83	84
85	86	87
88	89	90
91	92	93
94	95	96
97	98	99
100	101	102
103	104	105
106	107	108
109	110	111
112	113	114
115	116	117
118	119	120
121	122	123
124	125	126
127	128	129
130	131	132
133	134	135
136	137	138
139	140	141
142	143	144
145	146	147
148	149	150
151	152	153
154	155	156
157	158	159
160	161	162
163	164	165
166	167	168
169	170	171
172	173	174
175	176	177
178	179	180
181	182	183
184	185	186
187	188	189
190	191	192
193	194	195
196	197	198
199	200	201
202	203	204
205	206	207
208	209	210
211	212	213
214	215	216
217	218	219
220	221	222
223	224	225
226	227	228
229	230	231
232	233	234
235	236	237
238	239	240
241	242	243
244	245	246
247	248	249
250	251	252
253	254	255
256	257	258
259	260	261
262	263	264
265	266	267
268	269	270
271	272	273
274	275	276
277	278	279
280	281	282
283	284	285
286	287	288
289	290	291
292	293	294
295	296	297
298	299	300
301	302	303
304	305	306
307	308	309
310	311	312
313	314	315
316	317	318
319	320	321
322	323	324
325	326	327
328	329	330
331	332	333
334	335	336
337	338	339
340	341	342
343	344	345
346	347	348
349	350	351
352	353	354
355	356	357
358	359	360
361	362	363
364	365	366

```

QY 1380 GGGCTCCATGTCAGAGATGAATCTCTCAGGAGGCCAGACATATGATGAACCTCAGCCA 1439
Db 477 GGGCTCCATGTCAGAGATGAATCTCTCAGGAGGCCAGACATATGATGAACCTCAGCCA 418
QY 1440 TCCCAAGCTGGTTAAATTTCTATGGAGTGTGTTCAAGGAATACCCCATATACATAGTGCAC 1499
Db 417 TCCCAAGCTGGTTAAATTTCTATGGAGTGTGTTCAAGGAATACCCCATATACATAGTGCAC 358
QY 1500 TGAATATATAGCAATGCTGCTTCTGCTGAATACCTGAGGAGTCACGGAAGGACTTGA 1559
Db 357 TGAATATATAGCAATGCTGCTTCTGCTGAATACCTGAGGAGTCACGGAAGGACTTGA 298
QY 1560 ACCTTCCAGCTCTTGAAGATGCTGACGATGCTGTCGAAGCATGGCTTCTTGGAGAG 1619
Db 297 ACCTTCCAGCTCTTGAAGATGCTGACGATGCTGTCGAAGCATGGCTTCTTGGAGAG 238
QY 1620 TCACCAATTCATACACCGGACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1679
Db 237 TCACCAATTCATACACCGGACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 178
QY 1680 TGTGAAGATATCTGACCTTGAATGACAAAGTATGCTTGTGATGACCATGATGTCAGTTC 1739
Db 177 TGTGAAGATATCTGACCTTGAATGACAAAGTATGCTTGTGATGACCATGATGTCAGTTC 118
QY 1740 AGTCGGAAACAAAGTTTCCAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
Db 117 AGTCGGAAACAAAGTTTCCAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 58
QY 1800 CAGCAGCAAGTCAGACGCTATGGGATTTGG 1829
Db 57 CAGC-GCAAGTCAGACGCTATGGGATTTGG 29

RESULT 13
LOCUS AK080038
DEFINITION Mus musculus adult male aorta and vein cDNA, RIKEN full-length
          enriched library, clone:A330051M20 product:BMX non-receptor
          tyrosine kinase, full insert sequence.
ACCESSION AK080038.1 GI:26348342
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
TITLE Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
JOURNAL High-efficiency full-length cDNA cloning
MEDLINE Math. Enzymol. 303, 19-44 (1999)
PUBMED 99279253
REFERENCE 2
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL Prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 20499374
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
          Konno,H., Akiyama,J., Nishi,K., Kitzunai,T., Tashiro,H., Itoh,M.,
          Sumi,N., Ishii,Y., Nakamura,S., Hatzuma,M., Nishine,T., Harada,A.,
          Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
          Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
          Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
          Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
JOURNAL sequencing pipeline with 384 multicapillary sequencer
MEDLINE Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 20530913
PUBMED 11076861

```

---

```

REFERENCE
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
          Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
          Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
          Saito,T., Okazaki,Y., Gofobori,T., Bono,H., Kasukawa,T., Saito,R.,
          Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
          Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
          Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
          Quackenbush,J., Schriml,L.M., Staublei,F., Suzuki,R., Tomita,M.,
          Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
          Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
          Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
          Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
          Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
          Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
          Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
          Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
          Toyooka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L.,
          Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawayi,H., Kohsaki,S.,
          and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085650
PUBMED 11217851
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
          Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 21085650
PUBMED 11217851
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
JOURNAL Physical and Chemical Research (RIKEN), Laboratory for Genome
          Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
          RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
          Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp,
          URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
          Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site for further details.
          URL:http://genome.gsc.riken.go.jp/
          URL:http://fantom.gsc.riken.go.jp/
          Location/Qualifiers
            1..1014
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6J"
              /db_xref="FANTOM_DB:A530051M20"
              /db_xref="taxon:10090"
              /clone="A530051M20"
              /sex="male"
              /tissue type="aorta and vein"
              /clone_lib="RIKEN full-length enriched mouse cDNA library"
              /dev_stage="adult"
              145..>1014
              /note="unnamed protein product; BMX non-receptor tyrosine

```

CDS

```

kinase (MGDI:1101778, GB|NM_009759, evidence: BLASTN,
99%, match=2607)
/codon_start=1
/protein_id="BAC37811.1"
/db_xref="GI:26348343"
/translaton="MDNMESKSLLELLKKSQKKMSPNNYKERFLVLTSTLSY
YEYDKMRGSRKGSIEIKLRCEKVNLEBOTPEROYFPOIVYKGLLVYASNEES
RCOMLALOKIEIGNPHLLIKYHSGFVDFGKFLCCQOSCKAAPGCTLWEAYADLHAI
SDKHRAPEPERLLKIPRAVPVLIKMDASSGAILPOYDSYKSKSCSQSTSNRIYP
REDKPDWQVVRKUKSEBEDIACSNOLERNIASHSTSKMSWGFPESSSEEBENLHAYDW
FAGNISRSQSEQLL"
BASE COUNT      333 a      219 c      244 g      218 t
ORIGIN
Query Match      24.0%; Score 589; DB 11; Length 1014;
Best Local Similarity 79.9%; Pred. No. 6e-108;
Matches 766; Conservative 0; Mismatches 115; Indels 78; Gaps 3;
QY 7 ACGGAACAAGCTGAGAGCGGATGATATATGATACAAAATCTATTCTAGAGAACTCTT 66
DB 130 ACAGAGGAGCCCAATGGAGCAATATGAGAGTAAATCCATCTCGGAAGAACTGCTT 189
QY 67 CTCAAAAGATCAGACGAAAGAAAGAAATGTCACCAAAATATTAACAAGAACGCTTTT 126
DB 190 CTCAAAAGATCAGACGAAAGAAAGAAATGTCACCAAAATATTAACAAGAACGCTTTT 249
QY 127 GTTTTGACCAAAACAAACCTTTCTACTATGAATATGACAAAATGAAAAGGCGCAGAGA 186
DB 250 GTTCTTAACCAAAACAAAGCCTTTCTACTATGAATATGACAAAATGAAAAGGCGCAGAGA 309
QY 187 AAGGATCCATTGAAATTAAGAAATCAGATGCTGAGAGAAAGTAAATCTCGAGGAGCAG 246
DB 310 AAGGATCAATGAAATTAAGAAATCAGATGCTGAGAGAAAGTAAATCTCGAGGAGCAG 369
QY 247 ACGCCTGTAGAGAGACAGTACCATTTCAGATGCTCTATAAGATGGCTTCTCTATGTC 306
DB 370 ACACCGGTGAGAGACAGTACCATTTCAGATGCTCTATAAGATGGCTTCTCTATGTC 429
QY 307 TATGCATCAATGAAGAGAGCCGAGTCAGTGGTGTGAAGCATTTACAAAAGAGATAAGG 366
DB 430 TATGCATCAATGAAGAGAGCCGAGTCAGTGGTGTGAAGCATTTACAAAAGAGATAAGG 489
QY 367 GGTAAACCCACCTGCTGCTCAAGTACCATAGTGGTTCCTGCGAGGAGGAGTTCCTG 426
DB 490 GGAACCCCTCACCTGTTGATCAAGTATCAGTGGCTTCTGTTGGTGTGAAGTTCCTG 549
QY 427 TGTTCGACGAGAGCTGTAAGAGAGCCGAGGAGTACCCCTCTCGGGAAGCATATGCTAAT 486
DB 550 TGTTCGACGAGAGCTGTAAGAGAGCCGAGGAGTACCCCTCTCGGGAAGCATATGCTGAT 609
QY 487 CTCGATACCTGAGTCAATGAAGAGAAACACAGAGTTCACCTTCCAGACAGAGTCTG 546
DB 610 CTCGATACCTGAGTCAATGAAGAGAAACACAGAGTTCACCTTCCAGACAGAGTCTG 669
QY 547 AAGTACCTGGGAGTCTCTGCTCAAAATGATGACCACTTTCAGTACCACTCTA 606
DB 670 AAGATTCGAAGGCGAGTTCGCTTCTCAAAATGATGATCATCTTCAGGTGCCATCTA 729
QY 607 GCCCAATATGACACGAATCAAGAAAAAATATGGCTCCAGGCCACCATCTTCAAGTACC 666
DB 730 CCCCAATAT----- 738
QY 667 AGTCTAGCGAATATGACAGCACTCAAGAAAAATCTATGGCTCCCGAGCCAACTCAAC 726
DB 739 -----GACAGCTATTCAAGAAAAAGTTGTGTTCCAGGCCAACACAGCAAC 783
QY 727 ATGCAATATATTCAGGGAAGACTTCCCTGACTGCTGGGCAAGTAAAGAAACTGAAAAGT 786
DB 784 ATACGCTATATTCAGGGAAGACTTCCCTGACTGCTGGGCAAGTAAAGAAACTGAAAAGT 840
QY 787 AGCAGCAGCAGTGAAGATGTTGCAAGCAGTAAACCAAAAAAGAAAAAT---GTCAATCAC 843

```

```

841 -----AGCGAGGAAGACATTCGATGCAGTAACCACTGGAAGAAATATGCGCTCTCAC 894
QY 844 ACCACCTCAAGATTCATGGATTCCTCAGTCAAGTCTCATCTGAAGAGAGGAAAC 903
DB 895 AGCACCTCAAGATTCATGGATTCCTCAGTCAAGTCTCATCTGAAGAGAGGAAAT 954
QY 904 CTGATGATGATGATGCTGTTGCTGTAACATCCAGATCACAATCTGAACAGTACT 962
DB 955 CTGATGCTATGATGCTGTTGCTGTAACATCCAGATCACAATCTGAACAGTACT 1013

RESULT 14
BM542399      1029 bp      mRNA      linear      EST 20-FEB-2002
LOCUS      AGENCOURT_6436679 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539575
DEFINITION      5', mRNA sequence.
ACCESSION      BM542399
VERSION      BM542399.1 GI:18771874
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1029)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM12234 row: e column: 16
            High quality sequence stop: 493.

FEATURES             Location/Qualifiers
     source            1..1029
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5539575"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DHI0B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                     Average insert size 2 kb. Library constructed by Life
                     Technologies."
BASE COUNT      349 a      209 c      300 g      171 t
ORIGIN
Query Match      23.3%; Score 572.4; DB 12; Length 1029;
Best Local Similarity 93.6%; Pred. No. 1.3e-104;
Matches 597; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 832 AATGTGATCAGACACCTCAAGATTTTCATGGGAATTCCTGAGTCAAGTTCATCTGAA 891
DB 2 AACGTGAATCAGACACCTCAAGATTTTCATGGGAATTCCTGAGTCAAGTTCATCTGAA 61
QY 892 GAAGAGAAACCTGGATGATATGACGCTTGTGTTGATACATCTCAGATCAATCT 951
DB 62 GAAGAGAAACCTGGATGATATGACGCTTGTGTTGATACATCTCAGATCAATCT 121
QY 952 GAACAGTTACTCAGACAAAAGGAAAGAGAGAGCATTTATGGTTAGAAATTCGAGCAAA 1011
DB 122 GAACAGTTACTCAGACAAAAGGAAAGAGAGAGCATTTATGGTTAGAAATTCGAGCAAA 181
QY 1012 GTGGGAATGTACAGAGTCTCTCTATTTAGTAGGCTGTGAATGATAAAAAAGAACTGTC 1071
DB 182 GTGGGAATGTACAGAGTCTCTCTATTTAGTAGGCTGTGAATGATAAAAAAGAACTGTC 241

```

